VOLUME 36 SECOND SERIES NUMBER 7

OCT 9 1930

THE PHYSICAL REVIEW

A JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS

CONDUCTED BY

THE AMERICAN PHYSICAL SOCIETY

BOARD OF EDITORS

Managing Editor JOHN T. TATE, University of Minnesota, Minneapolis

I. S. BOWEN, R. S. MULLIKEN, H. D. SMYTH, E. U. CONDON, C. J. DAVISSON, F. K. RICHTMYER, K. F. HERZFELD, A. W. HULL, F. W. LOOMIS

OCTOBER 1, 1930

Use of the Refraction of X-Rays for the Determination of the Specific Charge of the Electron – – – – – – – – – – – – – – – – – – H. E. STAUSS 1

X-Ray Scattering Coefficient as a Function of Wave-Length and Atomic Number – ERROL M. CRESSY 7

Atomic Scattering Powers of Nickel, Copper and Iron for Various Wave-Lengths – RALPH W. G. WYCKOFF 17

Theory of Complex Spectra – – – – – – – – – – – – – E. U. CONDON 1

Zeeman Effect in the ZnH and CdH Bands – – – – – – – – WILLIAM W. WATSON 1

Principle of Identity and the Exclusion of Quantum States – – – – GILBERT N. LEWIS 1

On the Wentzel-Brillouin-Kramers Approximate Solution of the Wave Equation – L. A. YOUNG AND G. E. UHLENBECK 1

On the Origin of the Bands in the Spectrum of Mercury Vapor – – – – S. MROZOWSKI 1

Abnormal Shot Effect of Ions of Tungstous and Tungstic Oxide – – – JOHN S. DONAL, JR. 1

Distinction between Contact-Potential Effects and True Reflection Coefficients for Low-Velocity Electrons – – – – – – – – H. E. FARNSWORTH AND V. H. GOTTSCHALK 1

Current, Pressure, and Frequency Relationships for the Initiation and Maintenance of the Electrodeless Glow Discharge – – – – – – – – – – MILTON L. BRAUN 1

Absorption of Slow Hydrogen Positive Rays in Hydrogen – – – – ROBERT E. HOLZER 1

Velocity Distribution of Secondary Electrons from Molybdenum – – THEODORE SOLLER 1

Heat of Formation of Molecular Oxygen – – – – – – – – – L. COVELL COMBLEET 1

Dielectric Constant of Helium – – – – – – – – – – – – J. V. ATANASOFF 1

Potential and Potential Energy of Space Lattices – – – – – – – – C. N. WALL 1

Origin of the Variations in the Sun's Rotational Velocity – – – – – – ROSS GUNN 1

Triboluminescence – – – – – – – – – R. E. NYSWANDER AND BYRON E. CURRY 1

Letters to the Editor – 1

High Voltage Tubes, *M. A. Tuve, L. R. Hafstad and O. Dahl*—1283; Use of the Piezoelectric Interferometer for the Determination of Absorption in Gases for High Frequency Sound Waves, *Floyd Firestone and W. D. Hershberger*—1293; The Magnetic Moment of the Li Nucleus, *O. Stern and F. W. Bradmon*—1303; Wind Mixing and Diffusion in the Upper Atmosphere, *S. G. Brunt*—1305.

Book Reviews – 1

PUBLISHED BY THE AMERICAN PHYSICAL SOCIETY

[illegible] 450–456 ANNUAL [illegible]

THE AMERICAN PHYSICAL SOCIETY

OFFICERS OF THE SOCIETY

President:	HENRY G. GALE	University of Chicago, Chicago, Ill.
Vice-President:	W. F. G. SWANN	Bartol Research Foundation, Philadelphia, Pa.
Secretary:	W. L. SEVERINGHAUS	Columbia University, New York City
Treasurer:	G. B. PEGRAM	Columbia University, New York City
Editor:	JOHN T. TATE	University of Minnesota, Minneapolis, Minn.
Local Secretary for the Pacific Coast:	L. B. LOEB	University of California, Berkeley, Calif.

THE PHYSICAL REVIEW

Published by the American Physical Society at 450 Ahnaip St., Menasha, Wis. Executive office: University of Minnesota, Minneapolis, Minn. Two volumes annually, twelve issues per volume.

First series. From July 1, 1893, to January 1, 1913, the REVIEW was conducted by Edward L. Nichols, Ernest Merritt and Frederick Bedell, with the financial support of Cornell University. The first four volumes were issued annually; subsequent volumes, semi-annually; thirty-five volumes in all.

Second series. On January 1, 1913, the REVIEW was transferred to the American Physical Society as announced on the first page of the new series which was then begun. Since then it has been conducted by a Board of Editors elected by the Society. Semi-annual volumes begin in January and July. Since July 1, 1929, the REVIEW has been issued twice a month.

Contributors are referred to page 2 of the advertisements for further information.

Annual Subscription, throughout the Postal Union, $10.00 in advance.

Single copies of current volume, $.60.

General index, 1893-1920, $4.00, in advance.

Back Volumes, $6.00 per volume; Back Numbers, $1.00 per copy for issues prior to July 1, 1929 (Vols. 1-33), $.60 per copy for issues subsequent thereto. Postage is free if remittance accompanies order; otherwise postage is extra. No bound volumes can be supplied. The REVIEW offers to purchase a limited number of First Series, Vol. 1, No. 1 (at $2.00 each); Second Series, Vol. 7, No. 1, January 1916; Vol. 7, No. 2, February 1916; Vol. 7, No. 4, April 1916; and Vol. 7, No. 5, May 1916 (at $2.00 each).

Advertisements. Rates will be supplied by the Editor on request. Copy for each number must be received at Minneapolis two weeks prior to the dates of issue; old copy is used until new copy is received. Send electrotypes direct to the printers, George Banta Publishing Co., Menasha, Wis.

Discount of ten percent is allowed subscription agencies and members of the American Physical Society.

Changes of address and Complaints of failure to receive the REVIEW, in the case of members of the Physical Society, should be addressed to the Treasurer; in the case of other subscribers to 450 Ahnaip St., Menasha, Wis.; or to the Editor. New copies can be sent free in response to complaint of non-delivery only if notice of non-delivery is received within three months of date of issue.

Editorial Correspondence should be addressed to the PHYSICAL REVIEW, University of Minnesota, Minneapolis, Minn.

Entered at the postoffice at Menasha, Wis., as second class matter.
Acceptance for mailing at special rate of postage provided for in section 1103, Act of October 3, 1917, authorized September 6, 1926.

Second Series *October 1, 1930* *Vol. 36, No. 7*

THE

PHYSICAL REVIEW

THE USE OF THE REFRACTION OF X-RAYS FOR THE DETERMINATION OF THE SPECIFIC CHARGE OF THE ELECTRON

By H. E. Stauss*

Ryerson Physical Laboratory, The University of Chicago

(Received July 14, 1930)

Abstract

The discrepancies in the determination of the value of the fundamental constant e/m by the two different methods developed so far make desirable the development of new methods for its evaluation. The index of refraction of x-rays offers one method. According to the theories developed so far, in the case of a dispersive medium with no critical frequencies near the frequency of the incident radiation, the value of e/m may be expressed in terms of well-known constants, the wave-length of the incident radiation, and δ (one minus the index of refraction).

Index of refraction of quartz for MoKα_1 and Kβ radiation. A new method of using the prism with x-rays has been developed and used to determine δ. For crystalline quartz of density 2.6480 gm/cm^3, δ was determined for the $K\alpha_1$ radiation of molybdenum as $1.804 \pm 0.001 \times 10^{-6}$ and for the $K\beta$ as $1.436 \pm 0.001 \times 10^{-6}$.

The calculation of e/m is complicated by the discrepancies that have arisen in the absolute determinations of x-ray wave-lengths. Using the absolute wave-lengths as determined from the results of Bäcklin, Bearden, and Cork, the values of e/m* all lie between the values given by the spectroscopic and deflection methods, and they have nearly as great a range as the difference between the two older values. For this reason no definite conclusion can be made as to the most probable value of e/m.

Introduction

THE index of refraction of x-rays has, in addition to its intrinsic interest, an interest on account of the discrepancies in the evaluation of some of the fundamental constants existing at the present time. So far a number of different dispersion theories have been developed for the x-ray region, but they all agree in the limiting case when the frequency of the radiation is greatly different from the natural frequencies of the medium, when they all lead to

$$\delta = 1 - \mu = \frac{ne^2}{2\pi m\nu^2} \tag{1}$$

where μ is the index of refraction, n is the number of electrons per cm^3 in the refracting medium, e is the electronic charge, m the electronic mass, and ν the

* National Research Fellow.

frequency of the radiation. If this formula is assumed correct, one of the constants can be evaluated in terms of the others and δ. Ordinarily it would not seem feasible to determine any of the constants in terms of δ; but in view of the great discrepancy in the value of e/m as determined by the two methods developed so far,[1] there is need of additional methods of determining e/m. An attempt to evaluate this fundamental constant from the refraction of x-rays thus seemed desirable.[2]

In Eq. (1) n may be replaced by Ndz/M, where N is Avagadro's number, d the density of the medium, z the "molecular number", and M the molecular weight; and ν may be replaced by c/λ, where c is the velocity of light and λ the wave-length of the radiation. Ne/c in turn is Faraday's constant in e.m.u. With these values, Eq. (1) may be solved for e/m as

$$\frac{e}{m} = \frac{2\pi M\delta}{Fdz\lambda^2} \text{ in e.m.u.} \tag{2}$$

Of the quantities on the right, all except λ and δ are known or can be determined with high precision. The value of λ is in doubt on account of the discordant results obtained in its absolute determination by the several observers. Whenever λ can be assumed to be known, δ is the limiting factor in the evaluation of e/m. It will be noticed that Eq. (2) is independent of e, the value of which is also in doubt.

For the determination of δ it was decided to use the deviation produced by a prism and to develop the method suggested in an earlier paper.[3] There it was shown that the maximum deviation of a beam of x-rays can be obtained either with the beam striking the first face of the prism at nearly the critical angle of reflection, the method used by Larson, Siegbahn, and Waller,[4] or with the beam striking the second face internally at nearly zero glancing angle of incidence and leaving at nearly the critical angle of reflection. In the second method the refracted beam is less divirgent than the incident beam, and all wave-lengths suffer maximum deviation simultaneously. There is the disadvantage that an angular rotation must be used that must be measured with high precision.

Theory of Experiment

The experimental procedure adopted can be explained by reference to Figs. 1 and 2. Radiation from the slit S fell upon the prism P, which had its refracting edge A over the center of rotation of a spectrometer. Slightly to the front and to the side of the prism was a lead block B which acted with the prism as a sort of Seemann slit. The primary beam was defined by A and B. Refraction occurred in the neighborhood of A. After the primary and re-

[1] R. T. Birge: Phys. Rev. Supplement **1**, 43 (1929).

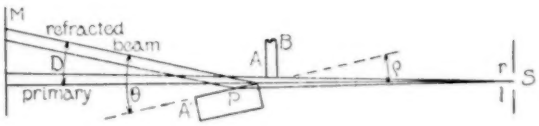
[2] It is a great pleasure to acknowledge that Professor A. H. Compton suggested the desirability of such a determination of e/m.

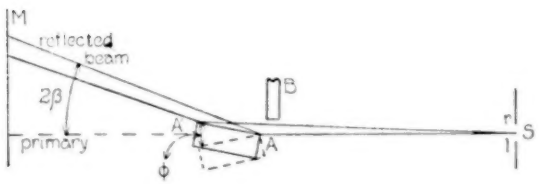
[3] H. E. Stauss, J. O. S. A. **19**, 167 (1929).

[4] A. Larson, M. Siegbahn, and T. Waller, Phys. Rev. **25**, 245 (1925).

fracted beams were registered on the photographic plate M, the prism was rotated through a small angle ϕ to the position of Fig. 2, and the beam was totally reflected from the face AA' at the angle β.

The value of δ is most readily calculated from $\mu = 1 - \delta = \cos \rho$, where ρ is the internal glancing angle with the face AA' and θ is the external angle. Assuming a single ray to be striking at A and assuming the ray to suffer no





Figs. 1 and 2. Diagrammatic sketch of experimental arrangement.

deviation at the first face, ρ and θ can be determined from ϕ, β, and D, where D is the angle of deviation.

$$\phi = \theta + \beta \quad \text{or} \quad \theta = \phi - \beta$$
$$\theta = D + \rho \quad \text{or} \quad \rho = \theta - D.$$

Substituting these values in the expression for μ, and solving for δ

$$\delta = D(D + 2\rho)/2. \tag{3}$$

The values of D and ρ need a slight correction (1/3″) on account of the deviation in passing through the first face; but this can easily be calculated.

In practice the rays from the slit S are divergent. The simplest experimental procedure is to make all measurements from the edges produced by the radiation from the edge r (of the slit S) that passed A. The errors introduced by measuring edges will be discussed later. In the case of the refracted beam, the ray defined by A and r lies within the beam. Theoretically refraction occurs all along AA'. Practically absorption limits the effective length of the refracting face. For the experiments reported here, the intensity of the refracted radiation that emerged 1 mm from A was only 9 percent of that which emerged at A. The change in ρ produced by the penetration produces a negligible effect upon θ; so the refracted radiation from r may be considered a band of parallel radiation, of width about 0.002 mm. A correction of this amount can be made to the position of the refracted beam. It is evident that the width of the refracted beam is determined almost solely by the slit S.

In practice it was found advisable to use two different widths of the slit S for the refracted beam and for the primary and reflected rays. The former

varied from 0.1 to 0.3 mm in different trials; the latter was always 0.012 mm. The changes in width could be made accurately with a micrometer screw that formed part of the slit. On account of the changes in width, the position in the refracted beam corresponding to r had to be calculated in order to obtain D_r. This can be done from the relation $\cos\theta = \mu\cos\rho$ or $\theta^2 = \delta + \rho^2$, since $D = \theta - \rho$. In practice the refracted beam was not as sharp as was to be expected; so to diminish the chances of error, both edges of the refracted beams were measured and averaged. In the case where two strong beams of different wave-length overlap, as is the case for the two $K\alpha$ lines, another correction is necessary, which can be shown to be approximately

$$\Delta D = \theta\Delta\lambda/\lambda.$$

The position of D_r is then obtained from the average position by the formula

$$D = D'_{Av} - \frac{\Delta\theta_l' - \Delta\theta_r'}{2} + \theta\frac{\Delta\lambda}{2\lambda} \tag{4}$$

where the primes apply to the greater slit-width, and $\Delta\theta_l'$ and $\Delta\theta_r'$ are the corrections to be made in θ (and hence in D) on account of the changes of ρ produced by changing the slit-width. No correction was made for the γ line overlapping the β line because of the great difference in intensity. The corrections were always less than 1 percent.

The accuracy of the experiment depends to a very great extent upon the second or refracting face of the prism. It should be good up to the very edge A because half of the radiation is transmitted within 0.28 mm of A. The prism used in this experiment was a piece of crystalline quartz which had been cut from a larger pitch-polished surface. It was good to a quarter of a fringe and showed no peculiarities at the edges. Nevertheless, the refracted beam was more diffuse than was to be expected, and there was a band of radiation lying between the primary and refracted beams whose presence had not been foreseen. If in the preparation of the plate, even a slight bending occurred very near the edge A, it might affect the results by changing the inclination of parts of the refracting face. Moreover, a polished surface is never a geometrical plane, and the exact character of the surface is problematical. The observed abnormalities may in some way be due to surface effects. For this work it has been assumed, though it is difficult to prove, that the cause of the diffuseness does not displace the centers of the refracted beams.

Another factor affecting the accuracy of the experiment is the location of the edges of the beams. All the beams have penumbras within which the intensity falls from the maximum to zero. The calculations made above have all assumed that measurements can be made to the edge of zero intensity. This obviously is incorrect. On the other hand, the grains in an x-ray plate have considerable size, and the blackening may even run past the zero edge. A priori it is difficult to predict the positions that will be measured. It is possible to design more involved arrangements that permit

the centers of lines to be measured, and the results obtained by the present method seem to justify further work with such an arrangement.

Experimental Arrangement and Procedure

All the experimental work was done in a basement room. The plate-holder and the slit S, Fig. 1, were two meters apart, with a Geneva optical spectrometer midway between them. The telescope and collimator of the instrument were removed; and a special turntable was made that rotated with the scale of the instrument, so that its motion could be measured. The table was provided with perpendicular slides for centering the refracting edge of the prism. The slit, taken from a Geneva double x-ray spectrometer, was aligned by taking advantage of the striations present in a beam of x-rays from a narrow slit, the analogues of the spiral formed by a pinhole. An optical prism with its base 90° to the refracting edges was mounted on the table, and the slit adjusted until one of the striations and an edge of the prism were parallel. Other objects mounted on the table thereafter could be aligned by reference to the striations. That the method was satisfactory is shown by the fact that no systematic deviations were ever observed in reflection or refraction which could be attributed to inclination of the slit to the prism edge or reflector.

The plate-holder was made approximately perpendicular to the x-ray beam in both the vertical and horizontal directions. As the apparatus was used, the deviation from normal was too small to produce appreciable differences in the results. To obtain the distance from the refracting edge to the photographic plate, a piece of glass was substituted for the latter, a steel rod pressed against it, and the distance from the end of the rod to the refracting edge of the prism measured with a travelling microscope. The length of the bar was obtained by comparison with a standard meter. It is believed that this distance was determined with an error of less than 0.1 mm or 0.01 percent.

The prism, for most of the trials, had its refracting edge formed by the polished surface and a narrow bevel at 45° to it. The theory of the prism shows that the value of the prism angle is of minor importance in producing the deviation and this angle was assumed to be $135° \pm 15°$. To test the validity of this procedure the bevel was removed and a face was cloth-polished at 90° to the refracting surface. Later the prism was beveled again because, in the polishing, the tool had overlapped and injured the edge slightly. This was shown by the diffuseness produced at the edge of the reflected beam.

The $K\alpha_1$ and $K\beta$ radiations of molybdenum were used for the measurements. A water-cooled, self-rectified tube was run at 42 K.V. maximum and 20 ma. The times of exposure were adjusted to avoid over-exposure of the different lines. Eastman x-ray plates were used and the measurements made with a Gaertner comparator reading to 0.001 mm.

The experimental procedure finally adopted was as follows. The prism was set in position for refraction, the slit S opened to the desired width, and an exposure made for 1 to 2 hours, with the position of the direct beam cov-

ered with a lead sheet. The slit was then narrowed to 0.012 mm, the lead removed, and the primary beam registered for about four minutes. This position was read on the scale of the spectrometer, thirty settings of the cross-hair in the micrometer eye-piece of one of the reading microscopes being made on the edge of one mark on the scale. The prism was then rotated until the face AA' (Fig. 2) could reflect, a matter of 7′, and an exposure made for about ten minutes, the lead sheet again blocking out the primary beam. This position was read thirty times on the same edge of the same mark on the scale. Thus the measurement of the angular displacement was thrown entirely upon the micrometer screw of the reading microscope. The screw was calibrated at three greatly different positions on the spectrometer scale. The order of the procedure was changed in as many ways as possible in the different trials in order to eliminate any systematic error due to the order. Measurements of the photographic plates were made to both the right and the left to avoid errors in estimating the edge in coming to it from within or without the beam. Nine positions were measured on each plate, and five measurements, to both right and left, made at each position.

Results

After the experimental method had been developed to a satisfactory state, eleven trials were made. In one of the trials the plate became fogged and difficult to measure. It is listed in Table I, but is not used in the final average. On another plate, where the value of ρ was 17″, the $K\beta$ line was not sufficiently strong to be measured accurately under the same conditions as the $K\alpha$ line. The first two trials were made on the original prism; the next four on the right-angled prism; and the rest on the final bevelled prism (all being the same block of quartz). The probable error of a single determination of δ, as deduced from the deviations of the various readings and the arbitrarily assigned errors was much less than the differences between the various trials. In the experiment every factor that seemed to have any influence on the

Table I. *Values of* δ.

Trial	ϕ	$S(mm)$	ρ	$\delta K\alpha_1 \times 10^6$	$\delta K\beta \times 10^6$
1	6′10″	0.322	48″	1.818	1.447
2	4′56″	0.322	59″	1.819	1.447
3*	6′ 9″	0.322	28″	1.756	1.457
4	5′40″	0.322	51″	1.821	1.450
5	4′44″	0.222	17″	1.791	
6	6′11″	0.222	1′40″	1.794	1.424
7	7′37″	0.322	42″	1.783	1.425
8	7′10″	0.222	2′41″	1.795	1.421
9	7′ 7″	0.322	2′22″	1.805	1.448
10	7′ 6″	0.222	2′20″	1.793	1.428
11	6′56″	0.122	2′14″	1.819	1.431
Average (omitting 3)				1.804 ± .001	1.436 ± .00085

* Plate fogged.

result was varied, except the slit-width for the reflected and primary beams. The values of δ show no systematic variation with any of the quantities. The variations seem to be due to the measurements of the plates. If it can be assumed that the variations in δ are purely accidental in origin, the probable error of the average may be calculated. The results of the trials are shown in Table I, together with the variables entering into the experiment. The character of the refraction spectrum is shown in Fig. 3.

At the time this investigation was begun, there seemed to be one absolute determination of x-ray wave-lengths of outstanding accuracy, and the evaluation of e/m seemed to be very feasible. Since that time a new set of experi-

Fig. 3. A refraction spectrum of the molybdenum radiation formed by quartz. From right to left are the reflected beam, the $K\alpha$ line, and the $K\beta$ line. At the left is the primary beam. Between it and the $K\beta$ line can be seen the radiation refracted at small angles.

ments has been performed with is not in agreement with the others and the question of the accuracy of the grating measurements of x-rays has been raised. The question of the value of x-ray wavelengths is now about as uncertain as the problem of the value of e/m. Consequently the value of e/m has been calculated for the values of λ as determined by crystal measurements and by Bäcklin,[5] Bearden,[6] and Cork,[7] without passing any judgment upon the relative accuracies of the determinations. The absolute values for the molybdenum radiation were calculated from the proportion

$$\frac{\lambda_1 \text{ (corrected)}}{\lambda_1 \text{ (crystal)}} = \frac{\lambda_2 \text{ (grating)}}{\lambda_2 \text{ (crystal)}}$$

where λ_2 is the wave-length used in the absolute determinations, and the values of λ (crystal) have been corrected for refraction. If it is assumed that the differences between the observers are due to the experimental arrangements and the gratings, for the present purpose it is sufficient to use only one wave-length of each observer to calculate the molybdenum lines. The results are shown in Table II. The lines used as standards were chosen because they were single.

To obtain the values of e/m, the most probable values of the constants as given by Birge[1] were used. The density of the quartz was determined as

[5] E. Bäcklin, Dissertation, Upsala (1928).
[6] J. A. Bearden, Proc. Nat. Acad. Sci. **15**, 528 (1929)
[7] J. M. Cork, Phys. Rev. **35**, 1456 (1930).

TABLE II. *Absolute values of* λ *as determined by different observers.*

Observer	λ used	$\lambda\alpha_1(AU)$	$\lambda\beta(AU)$
Crystal Method		0.7078	0.6314
Bäcklin	Mo $L\beta$	.7087	.6322
Bearden	Cu $K\beta$	.7095	.6329
Cork	Mo $L\beta$	.7100	.6333

2.6480 ± 0.0003 gm/cm^3; the molecular weight was taken as $60.06 \pm .03$. The results are listed in Table III. The probable errors are calculated from those

TABLE III. *The values of* e/m.

Observer or method	$e/m \times 10^{-7}$ determined from $K\alpha_1$	$e/m \times 10^{-7}$ determined from $K\beta$
Crystal	1.773	1.773
Deflection	$1.769 \pm .002$	
Bäcklin	$1.768 \pm .007$	$1.769 \pm .007$
Bearden	$1.764 \pm .001$	$1.765 \pm .001$
Cork	$1.762 \pm .0013$	$1.763 \pm .001$
Spectroscopic	$1.761 \pm .001$	

given by Bäcklin and Bearden without evaluation, and from an arbitrary error of 0.01 percent assigned to Cork's results. This is not meant to imply any judgment as to the relative accuracy of the experiments.

It is clear that the determination of e/m from the refraction of x-rays must wait upon the solution of the absolute values of x-ray wave-lengths. The range of values of e/m, as determined from the values of the different investigators, is nearly as great as the original differences in e/m. If the average of all the experimenters is taken, the resulting value[8], 1.765×10^7, lies midway between the two older values. The two later experiments lead to a value closer to the spectroscopic value, but this does not warrant a definite conclusion, since one of them gives a result nearly midway between the older values. Whenever the absolute values of x-ray wave-lengths are known accurately, the method of this paper offers a way of evaluating e/m that is independent of the older methods and of the same order of accuracy.

I wish to thank the physics department of the University of Chicago for the apparatus placed at my disposal, and Professor A. H. Compton for his interest in all stages of the work.

[8] If the dispersion formula had been used in the form

$$\delta = \sum_i \frac{n_i e^2}{2\pi m(\nu^2 - \nu_i)^2}$$

where ν_i is a critical absorption frequency of the dispersive medium, the values of δ and e/m would be increased by unity in the last significant place.

X-RAY SCATTERING COEFFICIENT AS A FUNCTION OF WAVE-LENGTH AND ATOMIC NUMBER

By Errol N. Coade
Ryerson Physical Laboratory, University of Chicago

Abstract

The mass scattering coefficients for Au, Sn, Ag, Fe, and C were compared with those of Al for wave-lengths between $\lambda = 0.48$A and $\lambda = 1.09$A by collecting the scattered radiation in an ionization chamber surrounding the scatterer on one side. The primary beam consisted of characteristic fluorescence radiation excited in radiators of Sn, Ag, Mo, and Se by a water-cooled tungsten tube. The ratios obtained are shown in the table.

With Hewlett's measurement of the absolute value of the mass scattering coefficient of carbon these ratios can be transformed into absolute scattering coefficients.

Introduction

When x-rays pass through an absorbing screen part of the energy of the beam is spent in ejecting photoelectrons and part is scattered. If μ is the total absorption we may write

$$\mu = \tau + \sigma$$

where τ is the true absorption due to the ejection of electrons, and σ the absorption due to scattering. The fraction of the energy scattered per gram of the scattering substance is σ/ρ, the mass scattering coefficient. The experiment to be described was undertaken to measure σ/ρ directly, and to aid in determining the value of the true absorption from the known values of the total absorption. Another purpose was to find reliable values of σ/ρ to check current theories of the scattering by complex atoms.

Direct measurements of the mass scattering coefficient of x-rays have been made by Barkla and Dunlop, Hewlett, Mertz and others. Barkla and Dunlop[1] observed the scattering at 90° of filtered white radiation from an x-ray tube striking scatterers of different materials. It may be pointed out that these experimenters did not measure the true scattering coefficient because all their measurements were made at right angles to the primary beam. This would be adequate if the angular distribution of scattering were the same for all elements; but such is not the case. We know that as heavier atoms are used the angular distribution changes, with an increasing preponderance of scattering taking place in the forward direction. This scattering in excess of that called for on Thomson's theory was at that time considered to be extra-radiation, not connected with the scattering process, and was purposely excluded from measurement. In their experiments the purity of the primary radiation left much to be desired, since it consisted of direct radiation from an x-ray tube filtered through absorbing screens. This was

[1] Barkla and Dunlop, Phil. Mag. **31**, 222 (1916).

used because with fluorescence radiation the intensity of the scattered beam was so small that it could not be measured accurately. Barkla and Sale[2] repeated the experiment with modifications.

For the same reason Hewlett[3] used direct radiation for his primary beam. This radiation was, however, from a molybdenum target, and was filtered through a zirconium screen, thus giving a large proportion of homogeneous rays. By moving his ionization chamber to measure the intensity of scattering at all azimuths he was able to integrate the intensity distribution and get the true scattering coefficient. His method, however, required so many observations that its application over a wide range of elements and wave-lengths would be very tedious.

It was undoubtedly for this reason that Mertz[4] adopted a large ionization chamber surrounding the scattering block, which would receive the scattering for all azimuths. His measurements, however, were confined to elements of low atomic numbers with Na the highest. He likewise used white radiation from an x-ray tube filtered through Al and Cu. The effective wave-length was varied by changing the tube potential. Mertz used solid scattering blocks and had troublesome corrections to make for the absorption of the blocks. His ionization chamber was filled with air and he had difficulty in obtaining sufficient intensity to extend his measurements to higher atomic numbers and longer wave-lengths.

Apparatus

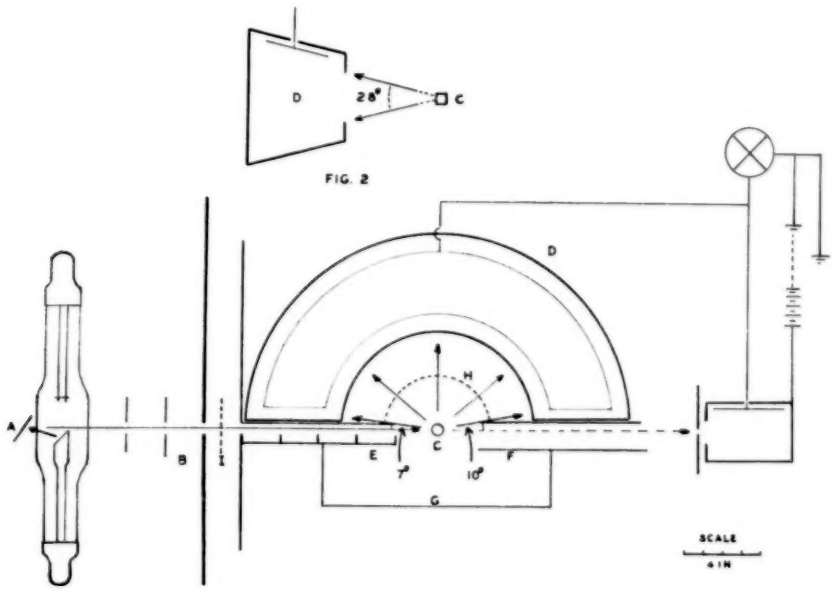
In the present investigation a modification of the chamber used by Mertz was employed to measure simultaneously the radiation scattered at all azimuths. Improvement was made in the quality of the x-ray beam by using fluorescence radiation excited in blocks of Se, Mo, Ag, and Sn by direct radiation from an x-ray tube. The scattering blocks were of Al, C, Fe, Ag and Sn. Suitable filters placed in the primary or secondary beam prevented fluorescence rays excited in the scatterer from entering the ionization chamber.

The source of x-rays was a water-cooled Siemens tube of the Coolidge type with a tungsten target surrounded by a copper screen. Current was supplied to the tube from a transformer, whose secondary current was rectified by a full-wave kenotron rectifier, using a balanced circuit. The tube potential was from 35 to 50 peak k.v. with a current of 35 ma.

The experimental arrangement is shown in Fig. 1. The radiators, A, were placed three cm from the x-ray tube and directly in front of a system of collimating screens, B, with openings one-half inch square. The beam of fluorescence radiation, which we will call the primary beam, passed through these openings and struck the scattering blocks at C. Since no direct radiation from the tube was allowed to reach the scatterer the beam consisted almost entirely of fluorescence K-rays.

[2] Barkla and Sale, Phil. Mag. **45,** 743 (1923).
[3] Hewlett, Phys. Rev. **20,** 688 (1922).
[4] Mertz, Phys. Rev. **28,** 891 (1926).

The ionization chamber, *D*, shown also in Fig. 2, was so shaped that it would receive the x-rays scattered between two planes meeting along the path of the primary beam at an angle of 28°. The chamber was of the same depth throughout, so that the path of all the scattered x-rays through it would be the same. The extreme rays entering the chamber from the center of the block made angles of 10° and 173° respectively, with the forward direction of the primary beam.



Figs. 1, 2.

The window of the ionization chamber was covered by celluloid. A protecting ledge about the window prevented x-rays from striking the walls or the electrode, which consisted of an aluminum plate, the same shape as the top of the chamber, and mounted in the upper portion of the chamber. The entire outside of the chamber and all the electrometer connections were covered with 1/32″ lead. Lead screens, *E* and *F*, were used to reduce the effect of air scattering from the path of the primary beam. A lead screen *G*, enclosing the scattering block and the window of the ionization chamber was found useful in excluding stray radiation. Absorbing screens were placed at *II* and *I* when required. The chamber was filled with argon at atmospheric pressure and was connected to a Compton electrometer, the sensitivity of which was about 10,000 mm per volt.

The scattering blocks were then mounted on supports of silk thread attached to wire frames and were so placed in the path of the x-ray beam that radiation from every part of the first slit reached every part of the scattering block.

Filters

Radiation of wave-length other than that desired was prevented from

entering the ionization chamber by filters of aluminum, silver, or tin placed in the primary or secondary beam or both. In addition, the potential of the tube was altered from 35 peak k.v. for Mo and Se to 50 peak k.v. for Ag and Sn. For scattering blocks of carbon and aluminum the K-radiation is so soft that it is absorbed by the air before reaching the ionization chamber. Consequently no filter was required. For the iron scatterer a filter of aluminum 0.2 mm in thickness was used to remove 99 percent of the K-radiation. When the scatterer was silver no filter was required for Mo radiation, but for Ag radiation filters of silver, each absorbing 50 percent of the K-rays, were placed in both the primary and secondary beams. The same filters were used for the scattering of Ag radiation by tin. When tin was used to scatter Sn rays a filter of tin was placed in the primary beam to absorb 50 percent of the characteristic radiation. This presumably absorbed more of the short wave-lengths present and prevented the excitation of fluorescence Sn rays. When Se rays were scattered by tin no filter was used, since the potential, 35 peak k.v., was too low to excite the fluorescence radiation. When gold was used to scatter Ag or Sn rays, an aluminum filter 1.2 mm thick was placed in the secondary beam to absorb the L-radiation.

The gold foil used was mounted on a celluloid base and was compared with an aluminum scattering block, likewise mounted on celluloid. In correcting for the scattering of the celluloid an amount was deducted equal to one-half of the scattering from a celluloid block of the same dimensions. For this and for all the other scattering blocks corrections were made for air scattering and for the natural leak of the electrometer system. Due to the length of air path in front of the chamber and to the size of the chamber itself, these corrections were appreciable, and their magnitude for longer wave-lengths and higher atomic numbers than those used made measurements by this method impracticable.

Measurement

The measurement of the mass scattering coefficient was made as follows: Let I_0 be the intensity when no scatterer is in the path of the primary beam, I_1 be the intensity when an aluminum block is in the path of the beam, I_2 be the intensity when an iron (or other) block is in the path, m_1 be the mass of the aluminum scatterer and m_2 be the mass of the iron (or other) scatterer. Then the ratio of the mass scattering coefficients of the two materials is

$$\frac{(\sigma/\rho)_{Fe}}{(\sigma/\rho)_{Al}} = \frac{m_1}{m_2} \times \frac{I_2 - I_0}{I_1 - I_0}.$$

The resulting ratios are shown in Table I and are plotted in Fig. 3. Data obtained by Barkla and Dunlop[1] are included for comparison.

The purity of the primary radiation was tested roughly by absorption methods and the beam was found to be approximately monochromatic. No attempt, however, was made to measure the wave-lengths accurately, because it was felt that better results could be obtained from published

TABLE I. *Ratio of mass scattering coefficients of elements to that of aluminum for different wave-lengths*

Scatterer	Radiator			
	Sn	Ag	Mo	Se
Au	5.49	6.75		
Sn	2.33	2.82	3.44	
Ag		2.65	3.51	
Fe	1.41	1.65	2.09	
C	1.07	0.95	0.87	0.79
Effective Wave-length (calc.)	0.48A	0.55	0.70	1.09

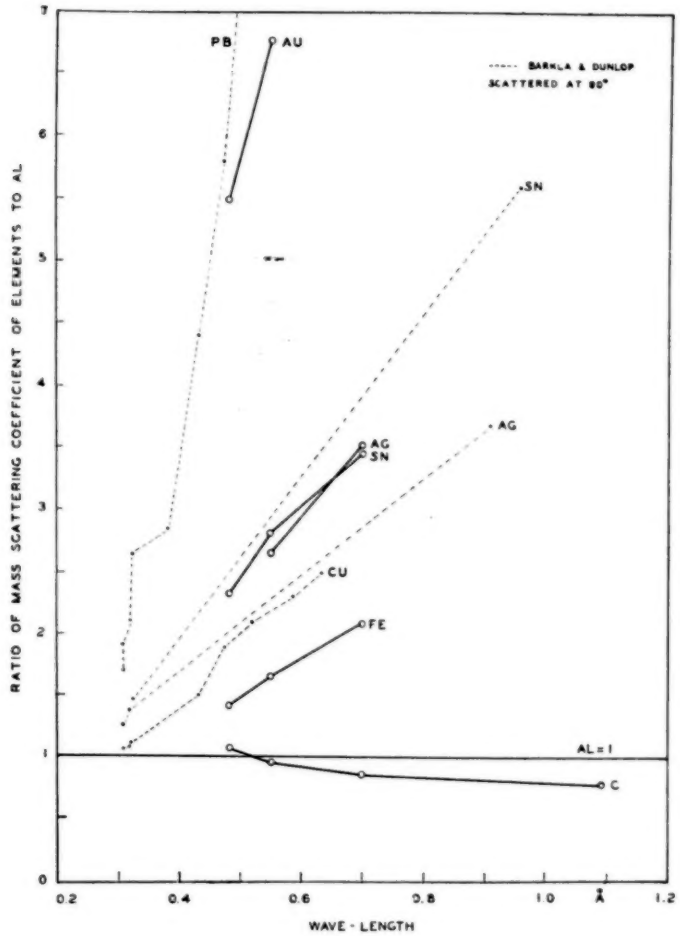


Fig. 3.

values for x-ray spectrum lines. Accordingly the wave-lengths of the K_α and K_β lines of each element were weighted according to the relative

intensities of each line,[5] and the average obtained. The weights are contained in the following table:

TABLE II.

Atomic No.	Element	Intensity ratio
34	Se	6.0
42	Mo	5.5
47	Ag	4.7
50	Sn	4.7

ABSOLUTE MASS SCATTERING COEFFICIENTS

The absolute values of the mass scattering coefficient can be deduced from Hewlett's measurements on carbon. He found[3] that $\sigma/\rho = 0.200$ for $\lambda = 0.7$A. The corresponding value of μ/ρ for the same wave-length was found also by Hewlett[6] to be $\mu/\rho = 0.532$. Substituting in

$$\frac{\tau}{\rho} = \frac{\mu}{\rho} - \frac{\sigma}{\rho}$$

one obtains $\tau/\rho = 0.332$

But τ is known to be very nearly proportional to λ^3 i.e.

$$\frac{\tau}{\rho} = K\lambda^3.$$

Therefore, for $\lambda = 0.7$A, $K = 0.968$ whence $\tau/\rho = 0.968\lambda^3$

From the equation

$$\frac{\sigma}{\rho} = \frac{\mu}{\rho} - \frac{\tau}{\rho}$$

We obtain the value of σ/ρ for carbon for the other wave-lengths as given in the following table:

TABLE III. *Mass scattering coefficients for carbon.*

Element	λ	σ/ρ (calculated)
Sn	0.48A	0.183
Ag	0.55	0.186
Mo	0.70	0.200
Se	1.09	0.330

With these values for carbon in Table I the absolute values of the mass scattering coefficients for the other elements were calculated and are shown in Fig. 4. Data obtained by Mertz for carbon are shown by the broken line for comparison.

[5] Siegbahn, Spectroscopy of X-rays, p. 96. These data are for the ratio of lines excited by cathode rays. That the ratio is the same for fluorescent lines has been shown in the case of Ag *K*-rays by A. H. Compton, Proc. Nat. Acad. Sci. **14,** 549 (1928).

[6] Hewlett, Phys. Rev. **17,** 284 (1921).

The writer is indebted to Professor A. H. Compton for suggesting this experiment and for invaluable advice during the course of the investigation.

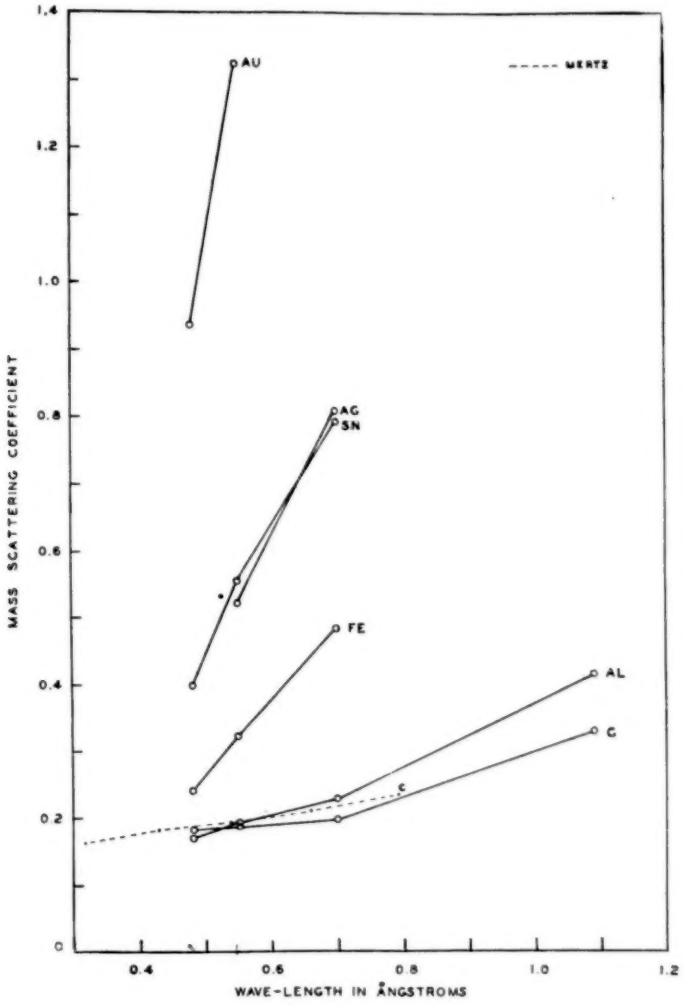


Fig. 4.

THE ATOMIC SCATTERING POWERS OF NICKEL, COPPER AND IRON FOR VARIOUS WAVE-LENGTHS

By Ralph W. G. Wyckoff
Rockefeller Institute for Medical Research, New York City

(Received June 26, 1930)

Abstract

Atomic scattering powers are measured for the atoms in metallic nickel, copper and iron when they are reflecting the $K\alpha$ lines of molybdenum, copper, nickel and iron. These data show that the scattering power of an atom passes through a minimum at its K absorption limit and attains a maximum at, or near, its "resonance" wave-length. The atomic F-curves of the nickel atom in metallic nickel and in NiO are practically identical.

THE data of this paper consist in measurements of the scattering powers of the metals nickel, copper and iron for the $K\alpha$ lines of molybdenum, copper, nickel and iron. Previous work from this laboratory with metallic copper and iron[1] showed that the reflecting powers of an atom might be considerably different for different wave-lengths. Studies[2] of NiO have indicated that the scattering power of nickel decreases as the wave-length used approaches its K critical absorption limit from the high frequency side and that it rises again after the limit is passed. The following results[3] suggest that this variation is generally to be expected.

F-curves for metallic nickel have been measured using the $K\alpha$ lines of molybdenum, copper, nickel and iron. The previous data from metallic copper and iron have also been extended through observations with copper, nickel and iron rays. The experimental procedures and the calculations leading to absolute F-values are the same[4] as those which have already been described. The $K\alpha$ lines of molybdenum were obtained by filtering the beam from a General Electric diffraction x-ray tube through a ZrO_2 filter; all other radiations were the unfiltered output of Siemens-Phoenix tubes equipped with appropriate targets. Calculation showed, however, that none of the measured $K\alpha$ reflections was, under the conditions of the experiment, contaminated by other wave-lengths.

The specimens studied were chemically pure, powdered metals which were passed through a fine meshed sieve and formed into suitable cakes by pressure applied in an ordinary vise. The surfaces of the preparations made in this way were filed and brushed before use in order to remove any super-

[1] A. H. Armstrong, Phys. Rev. **34**, 931 (1929).
[2] R. W. G. Wyckoff, Phys. Rev. **35**, 583 (1930).
[3] cf. R. W. G. Wyckoff, Phys. Rev. **35**, 215 (1930).
[4] A. H. Compton, X-rays and Electrons (New York, 1926), Chap. V; R. W. G. Wyckoff and A. H. Armstrong, Zeits. f. Krist. **72**, 319 (1929).

ficial layer which might show a preferred orientation. Two samples each of nickel and copper, one of domestic and the other of foreign manufacture, gave concordant results. The iron used was an imported preparation of pure electrolytic metal.

The relative intensities of the various measured reflections are listed in Tables I and II together with the intensities of (220) of NaCl which has served

TABLE I. *Relative intensities of the measured powder reflections from metallic nickel.*

Plane	Relative Intensities for Mo $K-\alpha$	Cu $K-\alpha$	Ni $K-\alpha$	Fe $K-\alpha$
111	$374._6$	$431._2$	$412._4$	$226._5$
200	$181._6$	$192._1$	$185._7$	[100]
220	[100]	[100]	[100]	—
113	$102._9$	$107._4$	—	—
222	34.1	32.0	—	—
400	—	—	—	—
133	30.6	—	—	—
240	28.6	—	—	—
224	18.3	—	—	—
333, 115	18.5	—	—	—
220 of NaCl	$571._7$	$107._7$	93.3	49.1

TABLE II.

Radiation	Reflection	Intensity of reflection (220, NaCl = 100)	F for 110, Fe
Cu $K-\alpha$	110, Fe	51.9	11.48
Ni $K-\alpha$	110, Fe	37.5	9.77
Fe $K-\alpha$	110, Fe	$609._2$	13.45
Cu $K-\alpha$	220, Cu	94.2	—
Ni $K-\alpha$	220, Cu	88.4	—
Ni $K-\alpha$	200, Cu	$167._7$	—
Fe $K-\alpha$	200, Cu	$153._7$	—
Fe $K-\alpha$	111, Cu	$330._1$	—

as standard. Since the absolute reflecting power of no crystal is known for any of the wave-lengths used except Mo $K\alpha$, it has, as before, been assumed that F(220, NaCl) is independent of wave-length and equal[5] to 15.62. Though this assumption is almost certainly not strictly accurate, it probably does not depart far from the truth. When absolute F's shall have been determined for these longer wave-lengths, it will be easy to shift the present results to conform with them.

The F-values calculated from the intensity data of Tables I and II are dependent upon the absorption coefficients that are used. Experimental values exist for most of these coefficients. Except for metallic copper when absorbing Cu $K\alpha$ and metallic iron when absorbing Ni $K\alpha$ radiations, they are in satisfactory agreement with the values calculated by Jönsson's general formula.[6] Even in the most unfavorable case, the choice of extreme values of absorption coefficients makes a difference of only 0.4 of an absolute F-unit.

[5] R. W. James and E. M. Firth, Proc. Roy. Soc. **117A,** 62 (1927).

[6] E. Jönsson, Uppsala Univers. Årsskrift (1928).

For the sake of uniformity and completeness, calculated μ/ρ's as shown in Table III, have been employed throughout.

The resultant reflection F's are recorded in Tables II, IV and V and in Figure 1. Several comparisons are possible between these data and the

TABLE III. *Absorption coefficients used in calculating F-values.*

Absorber	Density	Absorbed Wave-length Mo $K-\alpha$	Cu $K-\alpha$	Ni $K-\alpha$	Fe $K-\alpha$
Nickel	8.90	423	432	549	825
Copper	8.95	—	468	586	886
Iron	7.92	—	2556	3124	570
NaCl	2.16	18.0_1	$161._2$	$196._2$	$306._8$

TABLE IV. *Absolute F-values of the nickel atom in metallic nickel.*

Plane	sin θ/λ	Radiation Mo $K-\alpha$ (0.710 A)	Cu $K-\alpha$ (1.537 A)	Ni $K-\alpha$ (1.655 A)	Fe $K-\alpha$ (1.932 A)
111	0.246	18.06	14.94	15.92	16.05
200	.284	16.89	13.70	14.71	14.69
220	.402	12.91	10.57	11.29	—
113	.471	11.10	8.73	—	—
222	.492	11.64	8.32	—	—
400	—	—	—	—	—
133	.619	8.37	—	—	—
240	.635	8.36	—	—	—
224	.696	7.48	—	—	—
333, 115	.738	7.01	—	—	—

TABLE V. *Absolute F-values of the copper atom in metallic copper.*

Plane	sin θ/λ	Radiation Cu $K-\alpha$ (1.537 A)	Ni $K-\alpha$ (1.655 A)	Fe $K-\alpha$ (1.932 A)
111	0.241	16.60	—	14.59
200	.278	14.90	14.57	13.72
220	.392	11.53	11.18	—
113	.461	9.56	—	—
222	.481	9.10*	—	—

* The relative intensities of the reflections in this column (compared with 220) are those found by A. H. Armstrong (reference 1).

results of previous measurements. The $F(220, \text{Cu})$ for Cu $K\alpha$ is in fairly good agreement with existing information, but $F(110, \text{Fe})$ for the same radiation fails considerably below the published[7] value. Additional observations with iron powder from the source which supplied the earlier experiments make it seem probable that there was some preferential orientation in the previous powder cake. Of interest is the comparison that can be made between the F-curves of metallic nickel and of the nickel ion[8] in NiO. This comparison

[7] A. H. Armstrong, reference 1.

[8] R. W. G. Wyckoff, Phys. Rev. **35**, 583 (1930).

is shown graphically in Figure 2 for Mo $K\alpha$, Cu $K\alpha$ and Ni $K\alpha$ rays. The agreement is especially close for the Mo $K\alpha$ and Cu $K\alpha$ curves.

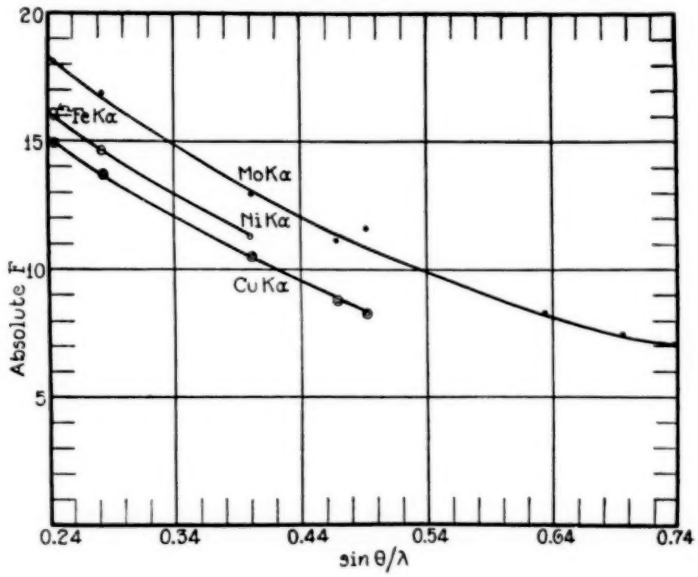


Fig. 1. The experimentally determined F-curves of metallic nickel for Mo $K\alpha$, Ni $K\alpha$ and Cu $K\alpha$ radiations. The two observed points for Fe $K\alpha$ rays are shown as large open circles.

Points on the F-curve for nickel with molybdenum radiation could be measured for reflections more complicated than (333; 115). Such reflections would be hard to determine with accuracy, however, because they are insignificant compared with the background of secondary nickel radiation

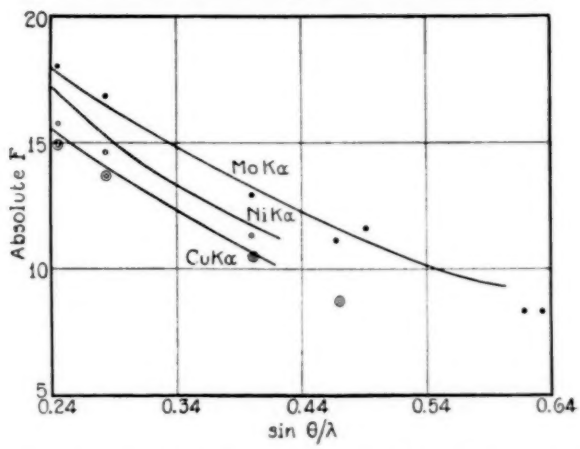


Fig. 2. The F-curves for the nickel atoms in NiO crystals are shown as full lines. Observed points for nickel atoms in metallic nickel appear as black circles (for Mo $K\alpha$), open circles (for Ni$K\alpha$) and ringed circles (for Cu $K\alpha$).

excited by the Mo $K\alpha$ lines. The first, (111), reflection of nickel occurs at so great an angle ($\sin\theta/\lambda = 0.246$) that much of the complete F-curve of metallic

nickel is of necessity extrapolated. This fact seems to prevent any important physical significance being attached to the electron-distribution curves that are calculated by Fourier analysis from the extrapolated forms of this *F*-curve.

The *F*-values of this paper show clearly the way in which the reflecting power of an atom varies with the wave-length of the x-rays scattered. The phenomena are illustrated by the (200) reflections of nickel as plotted in Figure 3. All of the other data of the present paper as well as those from NiO

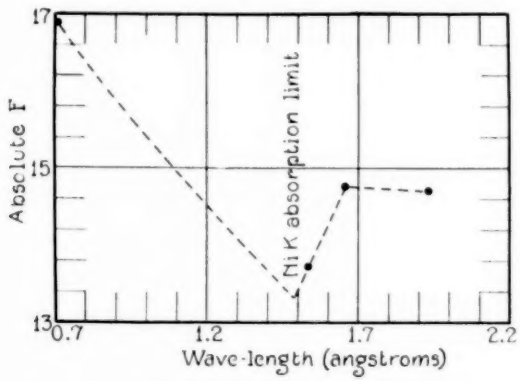


Fig. 3. The dotted curve indicates the variation of the scattering power of nickel with wave-length. The full circles are the measured values of the (200) reflection of metallic nickel.

agree with this figure in showing that the scattering power of an atom is at a minimum close to its *K* critical absorption limit, that it attains a maximum at, or near, its "resonance" wave-length, but that the reflecting power falls off only slowly on the long wave-length side of this frequency. The photographs made by Mark and Szilard[9] of RbBr using strontium and bromine radiations are obviously an expression of the same variation.

It is hoped through additional experiments to put these scattering powers for long wave-lengths upon an accurate absolute scale.

[9] H. Mark and L. Szilard, Zeits. f. Physik **33,** 688 (1925).

THE THEORY OF COMPLEX SPECTRA

By E. U. Condon

Palmer Physical Laboratory, Princeton University

(Received August 27, 1930)

Abstract

Extending the paper of Slater (Phys. Rev. **34,** 1293, 1929) on complex spectra, it is pointed out that assignment of definite electron configurations to spectral terms is an approximate procedure and only has meaning when the multiplet systems of the several configurations are widely separated. The effect of including spin terms is sketched. Non-diagonal matrix elements for the N-electron problem are reduced to corresponding elements for the 2-electron problem, as Slater did for the diagonal elements. Two-electron jumps occur because of the fact that spectral terms may not be precisely labelled by means of electron configurations.

IN A paper of this same name, Slater[1] has given a direct treatment of the application of the first order perturbation theory to the central field approximation to the atom model. He neglects spin forces in the Hamiltonian and so his results correspond to pure Russell-Saunders coupling of the angular momentum vectors, all intervals inside of the multiplets being zero. The electrostatic interaction gives the classification into multiplet levels and the calculations provide definite predictions concerning the intervals between the several multiplets which belong to the same electron configuration.

These intervals are found to be expressible in terms of certain double integrals over the radial factors of the eigen-functions of an electron moving in the central field which is made the starting-point of the perturbation calculation. For a configuration which gives n multiplets, there are thus $(n-1)$ intervals. Slater's first-order calculation expresses these $(n-1)$ intervals in terms of a fewer number (say, m) of integrals, usually, so even if one regards all of the integrals as independently adjustable, the theory predicts certain relations between the intervals. One may choose values for the m integrals as if they were independent so as to get the best fit possible in order to obtain a kind of test of Slater's results. Such a test of the results can be made comparatively simply without going into the more difficult question of determination of the best central field and the radial eigen-functions associated with it. If a good representation of the data is obtained on treating the m integrals as independent, the question still remains whether the m values assigned are compatible with the central field eigen-functions in view of the fact that they are not really independent. However, if a good representation cannot be obtained even by treating the integrals as independent, it certainly will not be improved when allowance is made for the fact that they are not independent

[1] Slater, Phys. Rev. **34,** 1293 (1929).

With such considerations in mind an attempt was made to apply Slater's results to a larger number of cases than he has treated in the examples at the end of his paper. It quickly became apparent that the intervals between the multiplets usually disagree badly with the first-order calculations. It is therefore necessary for an adequate theory of complex spectra to extend the calculations to a higher degree of approximation. Some results in that direction are the subject of this paper.

1. **Definitions and Notation.** The starting point, as with Slater, is a model of the atom in which N electrons each move, without influencing each other, in the same central field which has the potential energy, $-U(r)$. Only eigen-functions that are anti-symmetric in all pairs of electrons are used and so an eigen-function is specified by giving a *complete set* of $4N$ quantum numbers. A complete set consists of N *individual sets* which are called (n, l, m, k) these being the (n, l, m_l, m_s) of Slater. Each electron has four coordinates (x, y, z, s). The first three give its position and the fourth the z-component of spin angular momentum. For short the Greek letters, $\alpha, \beta, \gamma, \delta \cdots$ are written for separate individual sets. Also the capitals $A, B, C, D \cdots$ are written as abbreviations for different complete sets.

By Pauli's exclusion principle all individual sets in a complete set must be different, and two complete sets are not considered as different if they differ merely in regard to the order of listing of the same N individual sets. Nevertheless for definiteness a definite order of writing the individual sets in a complete set is adopted and adhered to during the calculations.

Since anti-symmetric eigen-functions are used there is no one-to-one correspondence between individual electrons and individual sets of quantum numbers. This means that an expression commonly used in spectroscopy such as "the excited electron is in a $4f$ state" refers to the presence of a $4f$ individual set in the complete set of quantum numbers. The electron configuration of a given complete set means the list of n, l values of the individual sets. Thus there are generally a number of different complete sets belonging to each configuration. Since the energy of a particle in a central field depends only on n and l, all of the complete sets belonging to a certain configuration have the same energy in the zeroth approximation from which the start is made.

For the one-electron eigen-function having the individual set $\propto$, written as a function of the first electron's coordinates, the notation $u_\alpha(1)$ is used, replacing Slater's $u(n_1/x_1)$. ψ is defined as

$$\psi = u_\alpha(1) \cdot u_\beta(2) \cdot u_\gamma(3) \cdots u_\xi(N)$$

so that the normalized eigen-function for the complete set A is

$$\Psi_A = (N!)^{-1/2} \sum_P (-1)^p P\psi$$

in which P stands for a permutation of the indices 1, 2, 3 $\cdots$ N in ψ relative to the $\alpha, \beta, \gamma \cdots \xi$. The summation extends over all $N!$ such permutations and p has the parity of P.

For the matrix component of any quantity as H which connects the states having the complete sets A and B, the Dirac notation $(A|H|B)$ is used, so that

$$(A \mid H \mid B) = \int \bar{\Psi}_A H \Psi_B$$

where $\int$ means integration over the $3N$ position coordinates and summation over the N spin coordinates.

2. **Formulation of the energy level problem.** The starting point is an exact solution of the quantum mechanical problem for a fictitious atom whose Hamiltonian is E where

$$E = \sum_1 \left[\frac{1}{2\mu}(p_{x1}^2 + p_{y1}^2 + p_{z1}^2) - U(r_1) \right] \tag{2.1}$$

where Σ_1 means that the same functional form is to be written down successively as depending on the coordinates of all N electrons and the results added together.

A form of the Hamiltonian for real atoms that is much nearer to the truth is

$$H = \sum_1 \left[\frac{1}{2\mu}(p_{x1}^2 + p_{y1}^2 + p_{z1}^2) - \frac{ze^2}{r_1} + V(r_1) M_1 \cdot s_1 \right] + \frac{e^2}{2} \sum_{1,2} \frac{1}{r_{12}} \tag{2.2}$$

In this the terms $V(r_1) M_1 \cdot s_1$ represent the energy of interaction of each electron's spin with its own orbital angular momentum. $V(r_1)$ is to be chosen in some way along lines of the semi-empirical discussions of the "screening for the spin doublets" of other workers. The question will not be discussed further in this paper. $M \cdot s$ is the scalar product of orbital and spin angular momentum for an electron. $\Sigma_{1,2}$ means a summation over all pairs of electrons, the two indices varying independently so that each pair is counted twice. Since the operators for E and H do not commute with each other, the matrix for H will not be diagonal in terms of the representation that is based on the eigen-functions of E. The problem of finding the energy levels for H is that of finding a transformation to the diagonal form for the matrix for H.

How the function $U(r)$ is to be chosen will not be discussed here. An approximate theoretical treatment, such as that of Thomas and Fermi, or a semi-empirical method of the sort studied especially by Hartree may be used. Slater studied the Hamiltonian (2.2) with omission of the spin term. Houston,[2] Bartlett,[3] and Gaunt,[4] have considered special cases of (2.1) counting the spin term and Goudsmit[5] has recently extended their results by a clever device.

[2] Houston, Phys. Rev. **33,** 297 (1929).
[3] Bartlett, Phys. Rev. **35,** 229 (1930).
[4] Gaunt, Proc. Roy. Soc. **A122,** 513 (1929).
[5] Goudsmit, Phys. Rev. **35,** 1325 (1930).

When the spin term is omitted the Hamiltonian H commutes with both the sum of the z-components of orbital angular momentum, and the sum of the z-components of spin angular momentum. Therefore H has no matrix components connecting states of E for which either of these sums is different. In terms of the quantum numbers introduced in §1 the quantities are Σm and Σk, the sums being over all the individual sets belonging to a complete set. Slater makes very good use of these results to calculate the energies of all the multiplets arising from a configuration (except for those configurations in which more than one multiplet of the same kind appears) without having to calculate any non-diagonal matrix components.

If the spin term is not omitted one still has the result that H commutes with the total sum of the z-components of both orbital and spin angular momentum although it no longer commutes with each sum separately. Therefore, even with spin counted there will be no matrix components of H connecting states for which the values of $\Sigma(m+k)$ differ.

There is another important property of the Hamiltonian which arises from its isotropy with regard to different orientations of the coordinate axes. The isotropy means that there is still a degeneracy, that of space quantization, associated with the Hamiltonian, so that the degeneracy of E is not completely removed by the inclusion of the spin and electrostatic repulsion terms which are the essence of the transformation from E a diagonal matrix to H a diagonal matrix. With each eigen-value of H can be associated a maximum value of $\Sigma(m+k)$ which is represented among the eigen-functions belonging to that eigen-value. This number is the quantum number J of that energy level. The isotropy then brings with it the result that there are other eigen-functions for which $\Sigma(m+k)$ has the values $J-1$, $J-2\cdots$, $-J$ all of which have the same eigen-value. The proof of this is best obtained by appeal to group theory.

If one uses the perturbation theory to find approximately the transformation from E diagonal to H diagonal, by treating $(H-E)$ as a perturbation, the success of the calculation in the first-order requires that the eigen-values of H which "grow out of" a particular eigen-value of E remain close together compared with the distance of the particular eigen-values of E from its nearest neighbor in the spectrum of E. This is a well known property of the perturbation theory and shows itself in many particular instances. Perhaps the Paschen-Back effect in the anomalous Zeeman effect is the best known of these. In it the effect of a uniform magnetic field on an atomic energy level is required. The first-order calculation is correct only if the spread of the energy levels growing out of the unperturbed level is small compared to the distance from the unperturbed level to its nearest neighbor in the unperturbed scheme. If this condition is not fulfilled then the second order perturbation becomes important. An important second-order correction implies that an important alteration of the eigen-function has taken place so that, when it is expanded in terms of the eigen-functions of E, it begins to have an important component of the eigen-function of the neigh-

boring level of E in addition to the eigen-function of the level from which it grew.

Therefore as the second-order correction becomes more important the quantum numbers which were appropriate to labelling the different eigenvalues of E become less and less appropriate for the labelling of the eigenvalues of H. They cease to be "good quantum numbers" to use a curiously apt expression introduced by Mulliken[6] in a discussion of the correlation of atomic energy levels with those of diatomic molecules.

3. Validity of configuration assignments. All discussions of complex spectra have hitherto been based on the idea that an electron configuration could be assigned uniquely to each energy level. That this procedure is only of approximate validity is seen at once from the foregoing discussion. The

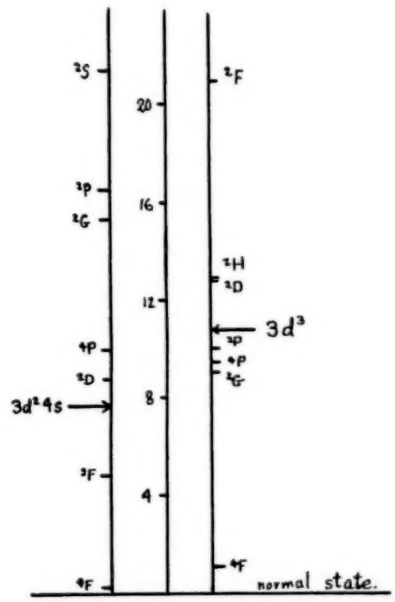


Fig. 1. *The two lowest configurations in Ti II.* Theory predicts another ^2D for $(3d)^3$ which has not been found; the arrow shows the center of gravity of the known terms, which is therefore not the exact center of gravity of the whole configuration. The center scale is in thousands of cm^{-1}.

criterion can be so formulated: *The assignment of a definite electron configuration to a group of multiplets is only exact insofar as the spread of the levels belonging to one configuration is small compared to the distance (on the energy level diagram) of the spread of levels belonging to the neighboring configurations.*

It is evident that this criterion for the case of the electron configuration quantum numbers is of the same form as the well known criterion for deciding between "weak" and "strong" magnetic fields in the Paschen-Back effect.

The importance of raising this point lies in the fact that complex spectra have already been analyzed in which the criterion for definite assignments

[6] Mulliken, Reviews of Mod. Phys. **2**, 60 (1930).

of electron configurations is not fulfilled. A noteworthy instance[7] is that of Ti II as analyzed by Russell.[8]

In Fig. 1 are plotted the levels corresponding to the two lowest configurations in Ti II as assigned by Russell. The arrows in each column give the centers of gravity of all the terms in the column, weights being assigned according to their values of $(2J+1)$.

It is evident that here, if anywhere, one may expect an appreciable effect of what may be suitably called interaction of neighboring electron configurations. Except for the overlapping of $d^3\,{}^2H$ and ${}^2D$ and of $d^3\,{}^2P$ and ${}^4P$ the distance between multiplets is large compared to intervals inside the multiplets, so that the criterion for Russell-Saunders coupling is fulfilled. The ${}^2H$ and ${}^2D$ can show no interaction, though, since they have no common value of J, whereas it is expected that the nearness of ${}^2P$ to ${}^4P$ will disturb the intervals. This is in fact the case, the ratio of the observed intervals being 3.82 against a theoretical (Landé) value of 1.67.

Perhaps the best way to present the situation is by reference to Slater's diagrams (loc. cit. p. 1301) of the energy matrix. Referring to his Fig. 2, one may suppose that the upper double-shaded square is the matrix for the different complete sets belonging to $(3d)^2 4s$ and the next double-shaded square is the similar matrix for complete sets belonging to $(3d)^3$. Slater says that the terms in the singly-shaded rectangles (these are the matrix components connecting the complete sets in $(3d)^2 4s$ with those in $(3d)^3$) are negligible. This is often the case but the point that is made here is that they are not negligible in a case where the distance between the centers of gravity of the terms coming from the same configuration is small or comparable with the spread of the multiplets arising from the two configurations separately. In case the configurations overlap, as they do in the special instance of Ti II one needs to consider the larger square which includes the two doubly-shaded squares and the two singly-shaded rectangles that border them.

If one next makes use of the exact theorem that there are no matrix components connecting complete sets for which $\Sigma(m+k)$ differs the large square is considerably simplified. It can be rearranged so that complete sets belonging to either configuration but having the same $\Sigma(m+k)$ are in adjacent rows and columns and then the large square corresponding to the two configurations taken together will break up into a series of smaller squares. If one wishes further to neglect the spin terms, as Slater does, then these smaller squares may be broken up into still smaller ones since then both Σm and Σk separately have to be equal in the quantum numbers labelling rows and columns.

The actual procedure in applying the diagonal sum method will be illustrated in terms of $d^2 s$ and d^3 although the numerical application of the formulas to $(3d)^2 4s$ and $(3d)^3$ in Ti II will be reserved to a later paper. One has first to set up the scheme of complete sets of quantum numbers for each config-

[7] I am indebted to my friend, Prof. J. E. Mack, for directing my attention to this case.
[8] Russell, Astrophys. J. **66**, 283 (1927).

uration, just as Slater does. For example, Table I gives the quantum numbers for the (*dds*) configuration.

TABLE I. *Sets of quantum numbers for (dds) configuration.*

$\Sigma k=$	3/2 (+++)	1/2 (+−+)	(++−)
$\Sigma m=4$		(220)	
3	(210)	(210) (120)	(210)
2	(200)	(200) (020) (110)	(200)
1	(2−10) (100)	(2−10) (−120) (010) (100)	(2−10) (100)
0	(2−20) (1−10)	(2−20) (−220) (1−10) (−110) (000)	(2−20) (1−10)

In this table are listed the m values that may be associated with the sets of spin (k) values that head the columns. The corresponding table for d^3 is Table II.

TABLE II. *Sets of quantum numbers for (ddd) configuration.*

$\Sigma k=$	3/2 (+++)	1/2 (++−)
$\Sigma m=6$		
5		(212)
4		(211) (202)
3	(210)	(210) (201) (2−12) (102)
2	(21−1)	(21−1) (200) (2−11) (2−22) (101) (1−12)
1	(21−2) (20−1)	(21−2) (20−1) (2−10) (2−21) (100) (1−11) (1−22) (0−12)
0	(20−2) (10−1)	(20−2) (2−1−1) (2−20) (10−1) (1−10) (1−21) (0−11) (0−22)

In these tables it is not necessary to list the quantum numbers for negative values of Σm and Σk as these do not give additional information.

The multiplet schemes corresponding to the two configurations are given superposed in the Table III, the first line in each cell being the contribution from d^2s and the second that from d^3.

TABLE III. *Multiplets for d^2s and d^3.*

$\Sigma k=$	3/2	1/2								
$\Sigma m=6$										
5		^2H								
4		^2H	^2G ^2G							
3	^4F ^4F	^2H	^2G ^5G	^2F ^2F					^4F ^4F	
2	^4F ^4F	^2H	^2G ^2G	^2F ^2F	^2D ^2D	^2D			^4F ^4F	
1	^4F ^4P ^4F ^4P	^2H	^2G ^2G	^2F ^2F	^2D ^2D	^2D	^2P ^2P		^4F ^4P ^4F ^4P	
0	^4F ^4P ^4F ^4P	^2H	^2G ^2G	^2F ^2F	^2D ^2D	^2D	^2P ^2P	^2S	^4F ^4P ^4F ^4P	

Now if the interaction between configurations can be neglected the diagonal sum method gives a value for the sum of the energies of all the multiplets

lying on each line separately in each cell. But if the interaction is important then the diagonal sums must be taken simply cell by cell, this gives only half as many equations and so the power of the diagonal sum method is greatly diminished (just as it is greatly diminished when spin is not neglected). Neglecting the interaction the diagonal sum method in this case is capable of giving the energy of each multiplet except that it can only give the arithmetic mean of the two ^2D's which arise from d^3. Allowing for the interaction it gives much less: now the ^2H and the ^2S are the only ones given directly, also given are the arithmetic means of the two ^4P's the two ^4F's, the three ^2D's, the two ^2P's, the two ^2F's and the two ^2G's.

In order to get the actual separation of the terms of similar multiplet character whose sums only are given one needs to have the non-diagonal elements of the energy matrix.

With these at hand one can set up the secular determinants for each cell and get the roots of the secular equations. This looks formidable at first sight because the determinants are of the order equal to the number of multiplets in each cell. But the diagonal sum method can be used to depress the order of the secular equations so that one needs only to solve a quadratic where there are two multiplets of the same kind, a cubic if there are three, and so on.

Perhaps it is helpful here to point out how the diagonal sum method works when spin is not neglected. In that case one can not write an equation that the sum of the terms in each cell of the Σm, Σk diagram is equal to the sum of the corresponding diagonal elements of the energy matrix. Also one needs a complete table covering the negative values of Σk as well as the positive. For simplicity the argument will be presented in terms of the sp configuration, which has already been fully treated by Houston.[9] One has for Σm,Σk:

sp

$\Sigma k=$	1 (++)	0 (+−) (−+)	−1 (−−)
$\Sigma m=1$	(01)	(01) (01)	(01)
0	(00)	(00) (00)	(00)
−1	(0–1)	(0–1) (0–1)	(0–1)

Arranging by values of $\Sigma(m+k)$ one sees that the values of $\Sigma(m+k)$ and the number of complete sets by which each value is realized is

$\Sigma(m+k)$	Realizations	Terms represented			
2	1	$^3P_2$			
1	3	$^3P_2$	$^3P_1$		$^1P_1$
0	4	$^3P_2$	$^3P_1$	$^3P_0$	$^1P_1$
−1	3	$^3P_2$	$^3P_1$		$^1P_1$
−2	1	$^3P_2$			

[9] Houston, Phys. Rev. **33**, 297 (1929).

The third column gives the energy levels whose sum is given by applying the diagonal sum method to all complete sets which have the same value of $\Sigma(m+k)$. Just as the diagonal sums, without spin, fails to give the separation between two multiplets of the same kind, so here with spin it fails to give the separation between two levels having the same J value. To get the separation one would need to solve a quadratic equation, which is what Houston did. And generally, the diagonal sums alone will give simply the arithmetic mean of all the terms in the configuration that have the same J value. To get them separately one has to use the non-diagonal elements and solve an algebraic equation whose degree is equal to the number of times the particular J value occurs. This statement is the starting point of Goudsmit's recent work on the transition from Russell-Saunders to jj-coupling.

Of course it is evident that if perturbations by an external magnetic field are included the method of diagonal sums gives the derivation of the g-sum rules that play such an important role in the theory of the anomalous Zeeman effect. But similarly if one wants to find the individual g's for n terms of the same J value then an algebraic equation of the n^{th} degree has to be solved, and this is true whether the n terms come from the same configuration or not.

4. Non-diagonal matrix elements. The diagonal sum method works only with the diagonal matrix elements of the energy and is quite powerful. It gives the energies of each multiplet belonging to a configuration if the interaction between configurations is neglected and if all multiplets are of a different kind. To go beyond this and find the separations between two multiplets of the same kind, or to allow for interaction of configurations and for other questions it is necessary to know the non-diagonal elements.

Two types of symmetrical function of the electron coordinates are of especial importance. One is of the form $F=\Sigma_1 f(1)$, that is, the sum of the same function of each one of the electron's coordinates occurring one at a time. The other is $G=\Sigma_{12} g(1,2)$, that is, the sum over all possible pairs of electrons, of a symmetrical function of the coordinates of both of them. Slater has carried out the calculation for the diagonal elements, the extension to the non-diagonal elements is very easy and almost exactly like Slater's work so only the results will be stated.

For a quantity of the type F, the matrix components, $(A\,|F|B)$, connecting states with the complete sets A and B are as follows:

(a) They vanish if B differs from A in regard to more than one individual set.

(b) If B differs from A solely in regard to one individual set then

$$(A\,|F|\,B) = \int_1 \bar{u}_\alpha(1) f(1) u_{\alpha'}(1)$$

where α and α' are the only individual sets of A and B respectively which are not equal.

(c) The diagonal element $(A\,|F|A)$ is worked out by Slater. Its value is

$$(A\,|\,F\,|\,A) = \sum_{\alpha} \int_1 \bar{u}_\alpha(1) f(1) u_\alpha(1)$$

where the sum extends over all the individual sets in the complete set A.

An immediate corollary of the results (a) and (b) is that if $f(1)$ is a quantity which is a diagonal matrix in the one-electron problem, then F is a diagonal matrix in the N-electron problem. This is perhaps the simplest way of proving that the total z-component of spin and the total z-component of orbital angular momentum are diagonal matrices in the representation used for this approach to the theory of complex spectra.

Similarly the matrix components, $(A\,|G\,|B)$, for a quantity of the type G can be reduced as follows:

(d) They vanish if B differs from A in regard to more than two individual sets.

(e) If B differs from A only in that two of its individual sets which one may call α', β' differ from the individual sets of A, called α, β, then

$$(A\,|\,G\,|\,B) = \Big[2 \int_{1,2} \bar{u}_\alpha(1) \bar{u}_\beta(2) g(1,2) u_{\alpha'}(1) u_{\beta'}(2) - 2 \int_{1,2} \bar{u}_\alpha(1) \bar{u}_\beta(2) g(1,2) u_{\beta'}(1) u_{\alpha'}(2) \Big].$$

(f) If B differs from A only in that its individual set, α', differs from the individual set, α, in A then

$$(A\,|\,G\,|\,B) = 2 \sum_{\beta} \Big[\int_{1,2} \bar{u}_\alpha(1) \bar{u}_\beta(2) g(1,2) u_{\alpha'}(1) u_\beta(2) - \int_{1,2} \bar{u}_\alpha(1) \bar{u}_\beta(2) g(1,2) u_\beta(1) u_{\alpha'}(2) \Big]$$

where Σ_β means a summation in which β runs over the $N-1$ individual sets that are common to A and B.

(g) The diagonal element $(A\,|G\,|A)$ is given by Slater. It is

$$(A\,|\,G\,|\,A) = \sum_{\alpha,\beta} \Big[\int_{1,2} \bar{u}_\alpha(1) \bar{u}_\beta(2) g(1,2) u_\alpha(1) u_\beta(2) - \int_{1,2} \bar{u}_\alpha(1) \bar{u}_\beta(2) g(1,2) u_\beta(1) u_\alpha(2) \Big]$$

the summation running over all pairs of individual sets.

These results show further simplifications for quantities that are independent of spin. For $u_\alpha(1)$ one has

$$u_\alpha(1) = v_\alpha(1)\, \delta(s_1, k_\alpha)$$

where $v_\alpha(1)$ is a function of the position coordinates only. In the preceding formulas $\int_1$ means integration over the position coordinates and summa-

tion over the spin coordinate. If $f(1)$ or $g(1,2)$ do not depend on spin the summation over s can be carried out at once. The results are:

$$\text{(b)}\quad (A\,|\,F\,|\,B) = \delta(k_a, k_{a'}) \int_1 \bar{v}_a(1) f(1) v_{a'}(1)$$

$$\text{(c)}\quad (A\,|\,F\,|\,A) = \sum_a \int_1 \bar{v}_a(1) f(1) v_a(1)$$

$$\text{(e)}\quad (A\,|\,G\,|\,B) = 2\Big[\delta(k_a, k_{a'})\delta(k_\beta, k_{\beta'}) \int_{1,2} \bar{v}_a(1)\bar{v}_\beta(2) g(1,2) v_{a'}(1) v_{\beta'}(2) - \delta(k_a, k_{\beta'})\delta(k_\beta, k_{a'}) \int_{1,2} \bar{v}_a(1)\bar{v}_\beta(2) g(1,2) v_{\beta'}(1) v_{a'}(2)\Big]$$

$$\text{(f)}\quad (A\,|\,G\,|\,B) = 2\delta(k_a, k_{a'}) \sum_\beta \Big[\int_{1,2} \bar{v}_a(1)\bar{v}_\beta(2) g(1,2) v_{a'}(1) v_\beta(2) - \int_{1,2} \bar{v}_a(1)\bar{v}_\beta(2) g(1,2) v_\beta(1) v_{a'}(2)\Big].$$

$$\text{(g)}\quad (A\,|\,G\,|\,A) = \sum_{a,\beta} \Big[\int_{1,2} \bar{v}_a(1)\bar{v}_\beta(2) g(1,2) v_a(1) v_\beta(2) - \delta(k_a, k_\beta) \int_{1,2} \bar{v}_a(1)\bar{v}_\beta(2) g(1,2) v_\beta(1) v_a(2)\Big].$$

A consequence of the limitation (d) is that there is no first-order inteaction between two neighboring configurations that differ in regard to more than two electrons. This result is hardly likely to be of much importanee for two configurations that differ in regard to as many as three electrons will, in general, lie in widely separated parts of the energy level diagram so the interaction would be negligible anyway.

This completes the reduction of the non-diagonal elements to integrations over the coordinates of only one or two electrons. If in particular $g(1,2)$ is of the form $1/r$ additional developments may be made in which $1/r_{12}$ is expanded in a series of spherical harmonics and so all the integrals can be expressed, as Slater does for the ones occurring in the diagonal elements, as a sum of certain integrals over the radial eigen-functions multiplied by coefficients that are certain integrals of spherical harmonics. The calculations involve some more general coefficients than Slater's a's and b's. Detailed developments of this part of the reduction will be reserved for a later paper.

5. Two-electron jumps. An immediate application of the results of the preceding section is to the question of "two-electron jumps." These are transitions between energy levels whose configurations differ in regard to two of the sets of nl values. Such transitions are usually weak compared to the more usual one-electron jumps. One may also have "zero-electron

jumps," that is, transitions between terms arising from the same configuration. This is the usual way of stating the case, in which one speaks as though each term is uniquely associated with a single configuration of the central field approximation.

The interaction between a light wave whose vector potential is given by $A(x, y, z)$ and an atom consisting of N electrons is measured by a term in the Hamiltonian of the form

$$\sum_1 A(x_1, y_1, z_1) \cdot p_1$$

where p_1 is the momentum of the first electron. This form embraces quadrupole and all multipole radiations. The dipole radiation which is the first approximation is obtained by replacing $A(x, y, z)$ by its value at the center of the atom and writing

$$A(0, 0, 0) \cdot \sum_1 p_1$$

which is valid if the wave-length is great compared to the size of the atom so that A does not vary much over the size of the atom. Now even the exact interaction is of the form of the sum of the same operator function of the coordinates of each electron summed over each of the electrons. Therefore one sees that its matrix components connecting complete sets which differ in regard to more than one individual set vanish.

In other words, if assignment of electron configurations to energy levels were an exact procedure there would be no two-electron jumps. Thus the existence of two electron jumps is an indication of a break-down of exact configuration assignments.

This point is similar to the one that the existence of inter-system combinations, i.e. transitions between states of different multiplicities, is an indication of the break down of exact assignment of L and S values of the terms, i.e. break-down in the Russell-Saunders coupling scheme.

There is an interesting question of language involved here. If the expression "electron jump" is to be translated into quantum mechanics as meaning a change in an (nl) individual set in going from an initial state to a final state, then, strictly, only one-electron jumps occur. But the exact eigen-function of each energy level has in its make-up components belonging to several different configurations of the central field model, whereas it has been the custom to assign to the energy level one configuration which is presumed to be the one that has the largest component in the expansion of the exact eigen-function. One-electron jumps are the only ones really occurring but since the exact eigen-function has other configurations in it than the principal one from which it derives its configuration name, there will appear to be two-electron jumps when the transitions are described solely by reference to the single configuration name that is assigned to each level by the customary procedure. It is the same with inter-system combinations: the selection rule $\Delta S = 0$ is exact, but the actual energy levels have components in their eigen-functions corresponding to more than one value of S; then when

one persists in labelling the terms simply with the value of S most strongly represented in the eigen-function he is confronted with apparent violations of this selection rule.

The break-down of exact configuration assignments also has important implications for the theory of relative intensity of spectral lines. Ornstein and Burger[10] have already shown that when inter-system combinations have appreciable intensity that they must be appropriately reckoned in the application of the intensity sum rules. This point is brought out clearly also in recent work by Harrison.[11] That is, the sum rules have to be applied to all the lines arising from transitions between all the terms of the initial and final configurations. Evidently all this has to be extended one step further for cases in which the interaction between two configurations is important. For simplicity suppose the initial state levels can be all given a fairly exact configuration assignment but that for the final state there are two configurations in interaction. Then the sum rules for intensities will have to be extended to summations over all lines terminating on any of the levels belonging to both of the interacting configurations. Detailed consideration of the intensity relations will be postponed to a later paper.

This paper was written at Stanford University during the summer quarter. The writer takes pleasure in this opportunity to thank the Stanford physicists for their cordiality and especially Professor G. R. Harrison for stimulating discussions on spectroscopy. He is also indebted to Professor J. E. Mack for discussions last winter at the University of Minnesota.

[10] Ornstein and Burger, Zeits. f. Physik **40**, 403 (1926).
[11] Harrison, J. Opt. Soc. Amer. **19**, 109 (1929).

THE ZEEMAN EFFECT IN THE ZnH AND CdH BANDS

By William W. Watson

Sloane Physics Laboratory, Yale University

(Received August 15, 1930)

Abstract

The Zeeman effect in the $^2\Pi \rightarrow {}^2\Sigma$ bands of ZnH at λ4326 and CdH at λ4509 which involve strictly case (a) $^2\Pi$ states and $^2\Sigma$ states with large ρ-type doubling is reported in detail. Field strengths up to 30,400 gauss were used, together with a large concave grating in a parallel light mounting which gives a dispersion in the third order of about 1.44A per mm in this region.

Analysis of the ZnH and CdH bands—Intensity relations among the 12 branches of each band are discussed. The ρ-type doubling in the $^2\Sigma$ states is found definitely to be proportional to $K+\frac{1}{2}$ rather than to K

The Zeeman effect patterns.—The positions and widths of the blocks of components for all the lines are computed by means of the quantum mechanics formulas of E. L. Hill. Observed and computed patterns are shown to be in excellent agreement on a number of details unique to these bands. Diagrams and tables of data showing these details are given for $H=16{,}400$ gauss for the lines of several branches.

Introduction

With the aid of the formulas recently developed by Hill[1] for the Zeeman effect in doublet molecular states, quantitative consideration of the rather complex Zeeman patterns in doublet bands is now possible. Agreements with the predictions of the theory have been obtained by Almy and Crawford[2] in the $^2\Pi \rightarrow {}^2\Sigma$ MgH bands, by the writer and W. Bender[3] for the lower rotational states in the $^2\Pi \rightarrow {}^2\Sigma$ CaH bands, and by Almy[4] for the $^2\Sigma \rightarrow \Pi^2$ OH bands. The present paper gives a comparison of the computed Zeeman pattern with some of the details of the effect of a magnetic field on the lines of the $^2\Pi \rightarrow {}^2\Sigma$ bands of ZnH at λ4326 and of CdH at λ4509 which differ from any patterns reported for these other bands.

A few measurements on the Zeeman effect in the λ4326 ZnH band have previously been given by Hulthén[5] and by the writer and B. Perkins.[6] This earlier work was confined to observations only on the most obvious features of the action of the field, notably the doubling of the lines of low rotational quantum number in the Q_1, R_1, and $^PQ_{12}$ branches[7] with component separation approaching at the origin $2\Delta\nu_n$, and ignored the obscurer details which

[1] E. L. Hill, Phys. Rev. **34,** 1507 (1929). Additional treatment by Hill of the effect in $^2\Sigma$ states when ρ-type doubling is not neglected is given by G. M. Almy, Phys. Rev. **35,** 1502 (1930).

[2] G. M. Almy and F. H. Crawford, Phys. Rev. **34,** 1517 (1929).

[3] W. W. Watson and W. Bender, Phys. Rev. **35,** 1513 (1930).

[4] G. M. Almy, Phys. Rev. **35,** 1495 (1930).

[5] E. Hulthén, Thesis, Lund 1923.

[6] W. W. Watson and B. Perkins, Jr., Phys. Rev. **30,** 592 (1927).

[7] Hulthén designated these branches Q_2, R_2, and Q_1 respectively. The notation used throughout the present paper follows the latest recommendations.

now are of major interest for comparison with the theory. Particularly puzzling was the reported absence of any effect on the lines of the ${}^{O}P_{12}$ and P_1 branches. The incompleteness of these data was due in part to the lack of any guiding theory, and apparently in part also to insufficient exposure-time when obtaining the spectrograms. New high dispersion plates were obviously necessary, and it was decided to include the CdH bands in the investigation in order to note the effect of the very near approach to strictly case (a) conditions in the ${}^{2}\Pi$ state (strong orientation of the electronic spin in the direction of the electric axis in the molecule). There is also a considerable increase in the magnitude of the ρ-type doubling of the lower ${}^{2}\Sigma$ state in the CdH bands, the effect of which is discussed below.

It should perhaps be emphasized that these band spectra due to diatomic hydride molecules (and the He_2 bands) constitute, because of the wide spacing of the rotational lines in their fine structure, the best available material for the complete study of the effects of a magnetic field. Even in these bands there is usually a certain amount of overlapping of field patterns from lines of different branches.

Experimental Procedure

As in the CaH investigation, the source was an intermittent d.c. arc operating in a modified Back chamber mounted between the pole faces of a large water-cooled Weiss electromagnet. The pole faces were 2 cm in diameter, spaced 6 mm apart, the maximum field strength in this gap being about 31,000 gauss. The moving electrode was a block of either Zn or Cd c.p. metal about 5 mm square and 3 to 4 mm thick riveted on to the flattened end of a length of 1/8 inch round brass rod. The fixed electrode was a thin strip of Cu, insulated from the pole face by a sheet of mica. A slow stream of hydrogen gas from a commercial cylinder was pumped at about 5 cm of Hg pressure through the arc chamber. The arc current was usually held at 2 amperes. Because of the relatively large size of the Zn or Cd electrode, about five hours of exposure could usually be obtained before renewal of the electrode became necesary.

The spectrograms were obtained in the third order of the large concave grating mounted in parallel light, the dispersion being from 1.43 to 1.46A per mm. Field strengths of 9,600, 16,400, and 30,400 gauss were used for both the ZnH and the CdH bands. The strength of the magnetic field was determined by measurements on the patterns of the Zn triplet λλ4680, 4722, and 4811 as well as that of the λ4413 Cd intercombination line (3/2 normal triplet). Exposure times of 9 and 10 hours sufficed for good registration of the ZnH bands on Eastman 40 plates at all field strengths, but the CdH band lines are more weakened when the field is thrown on, so that 18 hours of exposure were required with a field of 30,400 gauss to give their patterns with fair intensity. With the same setting without the magnetic field, good spectrograms of these bands could be obtained in 15 minutes. This no-field comparison and an iron arc spectrum were photographed on every plate in addition to the field exposure.

The Analysis of the ZnH and CdH Bands

These bands have been completely measured and analysed by Hulthén,[5,6], the latest interpretation being by Mulliken[9] to whose article reference should be made for an energy level diagram showing all the observed transitions. The 12 possible branches of $^2\Pi \rightarrow {}^2\Sigma$ bands arising from all the transitions obeying the selection rule $\Delta J = 0, \pm 1$ are present and with good intensity in both of these band systems. This is due to, or is to be correlated with, the fact that the upper $^2\Pi$ states are close to strict case (a), as judged by the large separation between the $^2\Pi_{\frac{1}{2}}$ and $^2\Pi_{1\frac{1}{2}}$ levels (330 cm^{-1} for ZnH, 1001 cm^{-1} for CdH), and that the doublet separations in the lower $^2\Sigma$ levels are quite appreciable. The lines of the satellite branches are thus nicely separated from those of the main branches and are of comparable intensity, which are points of importance for our present work. As a matter of fact, since among the satellite lines and lines for which $\Delta K = 2$ the intensity rule $Q > R > P$ holds as for the main branches, the lines of low and intermediate J values in the ${}^PQ_{12}$ and ${}^RQ_{21}$ satellite branches are just about as intense as the accompanying main branch lines (P_1 and R_2 respectively) in the ZnH bands,[10] and are even stronger than the main branch lines in the CdH bands.

In order to compute the Zeeman patterns, one must know the exact magnitude of the ρ-type doubling in the $^2\Sigma$ state. This doubling is measured directly by the separation of the satellite and main lines having the same K values. There is the possibility of obtaining four such $\Delta\nu$'s for each K in each band and the set of average values yields the constant a in the expression $\Delta\nu = a(K+\frac{1}{2})$ which has been shown[11] theoretically to be the correct formula. These data taken from the present measurements of the no-field comparison spectra, indicate definitely that the doubling is proportional to $K+\frac{1}{2}$ rather than to K, in agreement with the theory. The calculations give for the ρ-type doublet widths $0.223(K+\frac{1}{2})$ for ZnH and $0.592(K+\frac{1}{2})$ for CdH. Hulthén[12] gives $2.10(K+\frac{1}{2})$ for this same doubling in the corresponding lowest $^2\Sigma$ state with $v''=0$ for HgH. The course of this doubling in this sequence of similar molecules would seem to indicate the importance of the rotational distortion of the electronic motions as the principal cause.[11,3]

The Zeeman Effect in the ZnH Bands

The action of the magnetic field on the upper $^2\Pi$ state may be computed with Hill's quantum mechanics formulas,[1] it being sufficiently accurate to use his Eq. (12) to give the magnetic sublevels:

[8] E. Hulthén, Ziets. f. Physik **32**, 32 (1925).

[9] R. S. Mulliken, Phys. Rev. **32**, 388 (1928).

[10] It has been previously stated (ref. 9) that the satellite and $\Delta K = 2$ branches are, on the whole, only about half as intense as the main branch lines. Visual estimates of relative intensities in the present measurements indicate that for both ZnH and CdH the main Q branch lines are more intense than the ${}^PQ_{12}$ and ${}^RQ_{21}$ line more nearly in the ratio 10:8 or 10:9.

[11] J. H. VanVleck, Phys. Rev. **33**, 467 (1929).

[12] E. Hulthén, Zeits. f. Physik **50**, 319 (1928).

$$W = B\{(J+\tfrac{1}{2})^2 - \Lambda^2 + \beta(\Lambda^2+\tfrac{1}{2}) \pm \tfrac{1}{2}[(2J+1)^2(1-\beta)^2 + \Lambda^2(\lambda+\beta)(\lambda-4+5\beta)]^{1/2}\} \quad (1)$$

The upper sign is to be associated with the F_2 levels ($J=K-\frac{1}{2}$), and the lower sign with the F_1 levels ($J=K+\frac{1}{2}$). The quantity $\beta=(\Delta\nu_n/B)M/J(J+1)$, $\Lambda=1$ for a Π state, and for this (0,0) ZnH band $B=7.3$ cm^{-1} and $\lambda=+45$cm^{-1}. Calculation shows that the $2J+1$ sublevels formed from each original level of total angular momentum J are symmetrically placed about the no-field position.

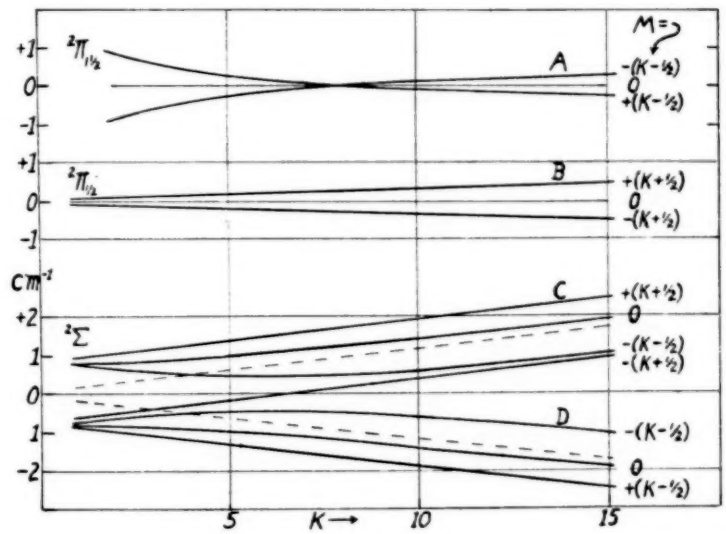


Fig. 1. Magnetic energy levels for the $^2\Pi$ and $^2\Sigma$ states of the ZnH λ4326 band. Only the levels for the extreme values of M are shown; the other magnetic levels should be spaced quite uniformly between these extreme levels. Note that the displacements in the $^2\Pi$ states are symmetrical about the no-field position. The dashed lines for the $^2\Sigma$ state represent the no-field positions of the two members of the ρ-type doublet. The positions and widths of the blocks of components for the lines of all the branches may be computed from the displacements in this diagram, noting that $\Delta M=0$, and ±1 only, and taking the proper ΔK for the branch in question.

The behavior of the $^2\Sigma$ state in the field is discussed by Almy.[4] If $\Delta\nu_n$ for the field strength in question is considerably larger than the ρ-type doubling at the origin of the band, there should be a molecular Paschen-Back effect, with the spin oriented either parallel or anti-parallel to the field. But the internal field which tends to orient the spin is still present, and therefore each member of this twice normal doublet is split into $2J+1$ magnetic sublevels. As K increases the spread of these magnetic levels about the no-field position should be from about $+\Delta\nu_n$ to $-\Delta\nu_n$, representing the various space orientations of the spin in the field. The position of these sublevels is given quantitatively by Hill's formula:

$$W = \pm\tfrac{1}{2}\{\nu_0^2 + 4M\nu_0\Delta\nu_n/(K+\tfrac{1}{2}) + 4\Delta\nu_n^2\}^{1/2} \quad (2)$$

where ν_0 is the ρ-type doublet width for the K level being considered, the $+$ sign is for the F_1 levels ($J=K+\frac{1}{2}$), and the $-$ sign for the F_2 levels ($J=K-\frac{1}{2}$). For the two outermost levels $M=\pm(K+\frac{1}{2})$ of the $J=K+\frac{1}{2}$ states, the energies are given by

$$W = \nu_0/2 \pm \Delta\nu_n. \tag{3}$$

For the sake of brevity, some of the principal features of the calculated and observed Zeeman patterns for one field strength only, 16,400 gauss, will be given here. All of the conclusions are substantiated by the measurements at the lower and higher field strengths. Fig. 1 is a plot to scale showing the course of the magnetic levels for which $M=\pm J$ for the $^2\Pi$ and the $^2\Sigma$ states when $\Delta\nu_n=0.77$ cm^{-1} (16,400 gauss). The levels for all of the intermediate values of M are spaced quite uniformly between these extreme levels. From this level diagram, paying due regard to the selection rule $\Delta M=0,\pm1$ and

TABLE I. *Comparison of calculated and observed patterns for $Q_1+{}^QR_{12}$ branches. ZnH (0,0) band. $H=16{,}400$ gauss. These results are plotted in Fig. 2. All calculations and observations in cm^{-1} with reference to the no-field position of Q_1 as origin. $+$ values are on the high frequency side, $-$ values on low frequency side. * indicates inaccurate because of an overlapping pattern.*

K''	Q_1 block		${}^QR_{12}$		Cross-over components	
	Obs.	Calc.	Obs.	Calc.	Obs.	Calc.
0	−0.80	−0.77	+0.67	+0.77		
1	−0.76	−0.63	+0.79	+0.72 to +1.08		
2	−0.63	−0.55	+0.72 to +1.24	+0.70 to +1.32	with ${}^QR_{12}$	
3	−0.60	−0.49	+0.55 to +1.60	+0.60 to +1.60	blocks	
4	−0.60	−0.15 to −0.63	+0.88 to +1.78	+0.82 to +1.85	+0.65	+0.62
5	−0.67 to −0.10	−0.60 to −0.00	+1.09*to+2.06	+0.88 to +2.10	+0.61	+0.60
6	−0.50 to +0.11	−0.58 to +0.10	+1.15 to +2.52	+0.95 to +2.35	+0.52	+0.55
7	−0.54 to *	−0.55 to +0.15	+0.95 to +2.80	+1.05 to +2.60		
8	−0.49 to +0.61	−0.52 to +0.50	+1.14 to *	+1.16 to +2.85	with Q_1	
9	−0.45 to +0.47	−0.50 to +0.48	+1.37 to +3.20	+1.30 to +3.15	blocks	
10	−0.46 to +0.50	−0.45 to +0.47	+1.70 to +3.38	+1.46 to +3.38		
11	−0.34 to +0.43	−0.40 to +0.42	+1.80 to +3.60	+1.60 to +3.62		

to the proper changes in K for the various transitions, the blocks of components in the Zeeman patterns may be drawn. Comparison with the similar energy level diagram given by Almy for OH (reference 4, Fig. 4) is of interest. The arrangement of the levels in the $^2\Sigma$ state is almost identical because of the approximate equality of the ρ-type doubling, but in the $^2\Pi$ state there is a marked difference, noteworthy being the crossing of the levels at $K=9$ in the $^2\Pi_{1\frac{1}{2}}$ set and the very slow, gradual increase in the spread of the levels in the $^2\Pi_{\frac{1}{2}}$ set.

We at once have the explanation of Hulthén's findings that the lines in some of the branches become doublets in the field, with doublet interval approaching $2\Delta\nu_n$ at the origin. These branches all represent transitions from the $^2\Pi_{\frac{1}{2}}$ upper level, and it is seen that for low K values the magnetic levels are here closely spaced about the no-field position. The magnetic sublevels in the $^2\Sigma$ state tend to group themselves into two sets separated by an inter-

val of about $2\Delta\nu_n$, and hence the possible transitions from the $^2\Pi_{\frac{1}{2}}$ levels to the levels in these two sets will be observed as two blocks of unresolved lines with this separation. Some examples of these field doublets may be seen in Fig. 5.

The only field doublet with really sharp components observed on our plates is that for the $Q_1(\frac{1}{2})$ line. This happens to be the only transition from the anomalous lowest $^2\Pi_{\frac{1}{2}}$ level with $K=1$, $J=\frac{1}{2}$, which is magnetically insensitive,[13] to the lowest $^2\Sigma$ level with $K=0$, $J=\frac{1}{2}$, which splits into a twice normal doublet. The observed doublet interval for this line is 1.47 cm^{-1}, as against $2\Delta\nu_n=1.54$ cm^{-1}.

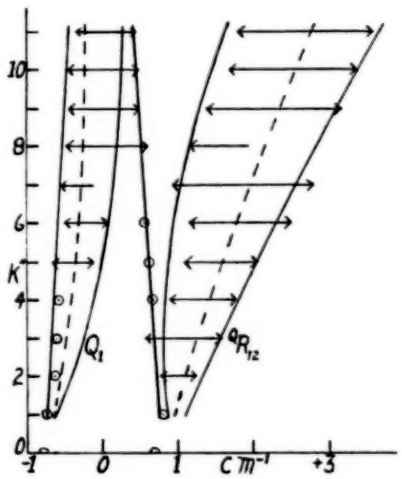


Fig. 2. Comparison of the computed and observed Zeeman patterns for the $Q_1+{}^QR_{12}$ lines of the ZnH band at 16,400 gauss. All displacements are reckoned with respect to the no-field Q_1 line positions. The dashed lines represent $M'=M''=0$, and divide equally the total number of components in each block. The crossing-over lines between the two blocks represent the two possible transitions to the $M=-(K+\frac{1}{2})$ level of the $^2\Sigma$ state in Fig. 1. Observed narrow components are represented by small circles, broader components by double-headed arrows.

Table I and Fig. 2 give a detailed comparison of the positions and widths of the blocks of Zeeman components, as obtained from Fig. 1, with the experimental values for the $Q_1+{}^QR_{12}$ lines. Such main lines and satellites must be considered together for each K'', for these transitions become intermingled in the field. The agreement in general is seen to be good. An interesting detail is the detection of the component crossing over from its association with the broader $^QR_{12}$ block to the narrower Q_1 block as K increases. This component arises from two transitions to the $M=-(K+\frac{1}{2})$ level which is to be noted in Fig. 1 crossing over between the two groups of magnetic levels in the $^2\Sigma$ state. The 30,400 gauss plate shows this narrow crossing-over component much more definitely both in these patterns as well as in those of other lines, particularly in the $^SR_{21}$ patterns. The low frequency edges of the narrower Q_1 portions of these patterns are stronger because of the greater

[13] Cf. Hill, reference 1 page 1516.

crowding of the components on this side (the dashed lines in the diagrams of Zeeman patterns represent $M'=M''=0$, and divide equally the total number of components in the block in each case). The greater width of the blocks of components for the lines of the satellite branches, such as the positive or ${}^{Q}R_{12}$ blocks of Fig. 2, and the ${}^{O}P_{12}$ and ${}^{S}R_{21}$ branches, is caused by the reversal of the order of the magnetic levels in the transitions $A\rightarrow C$ or $B\rightarrow D$ in Fig. 1. Remembering that M can change by only 0 or ± 1 in the transition, it is seen

TABLE II. *Zeeman patterns of the first ${}^{O}P_{12}$ lines. ZnII, (0,0) band. $H=16{,}400$ gauss. Displacements with respect to no-field position.*

J''	Weak neg. component Obs.	Calc.	Strong pos. component Obs.	Calc.
$1\frac{1}{2}$			+0.50	+0.60
$2\frac{1}{2}$	−1.30	−1.20	+0.46	+0.51
$3\frac{1}{2}$	−1.41	−1.40	+0.06 to +0.65	−0.03 to +0.80
$4\frac{1}{2}$	−1.60	−1.60	−0.08 to +0.75	−0.21 to +0.83
$5\frac{1}{2}$	−1.82	−1.82	−0.14 to +0.76	−0.40 to +0.85
$6\frac{1}{2}$	−1.97	−2.00	−0.33 to +0.74	−0.50 to +0.90

that the energy shifts are greater for these transitions than for the transitions $A\rightarrow D$ and $B\rightarrow C$. This is not the case for the CdH band patterns (see below). Fig. 5 shows several of these $Q_1+{}^{Q}R_{12}$ patterns.

The ${}^{O}P_{12}$ and P_1 branches are now found to exhibit Zeeman patterns quite similar to all of the other branches. The ${}^{O}P_{12}$ branch proceeds to longer wave-lengths from the ${}^{2}\Pi_{\frac{1}{2}}$ origin in a region clear of all other branches. These lines have no satellites, and yet, in agreement with the theory, they

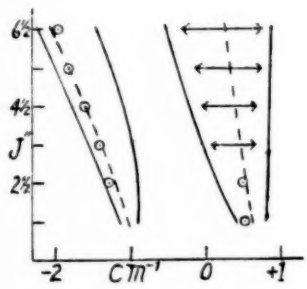


Fig. 3. Computed and observed patterns for the first lines in the ${}^{O}P_{12}$ branch of the ZnH band with $H=16{,}400$ gauss. The negative block is observed only as a weak, rather narrow line, the position of which in every case agrees well with the portion of the predicted block in which the components are more crowded together.

show two blocks of components coming from the transitions from the ${}^{2}\Pi_{\frac{1}{2}}$ magnetic levels to the two groups of magnetic levels in the ${}^{2}\Sigma$ state in Fig. 1. The comparison of the predicted and observed patterns for the first lines of the branch is shown in Fig. 3 and in Table II. In agreement with the theory, no crossing-over field component is found, since because of the selection rule for M and the fact that $\Delta K=2$ for these lines, transitions to the crossing-over magnetic level in the ${}^{2}\Sigma$ state cannot occur. The negative block is observed only as a weak narrow line whose position agrees well with that

of the center of the predicted components. Even though the positive block is considerably the broader, it is much stronger than the narrow negative block; about in the ratio 10:1 for ${}^{O}P_{12}(6\frac{1}{2})$. The same great intensity difference for the two blocks of a pattern is observed for the lines of the R_1 and ${}^{S}R_{21}$ branches, which also have no satellites in the no-field spectrum. This result is apparently contrary to the observations of Almy[4] on corresponding lines in the OH bands, and must be related to the much closer approach of ZnH to case (a) in the ${}^2\Pi$ state. For in the CdH bands, for which the ${}^2\Pi$ state is even closer to case (a), and with a larger doubling in the ${}^2\Sigma$ state, there is no trace of these weaker blocks in the patterns of these three branches. These transitions in the absence of an external field would involve the ex-

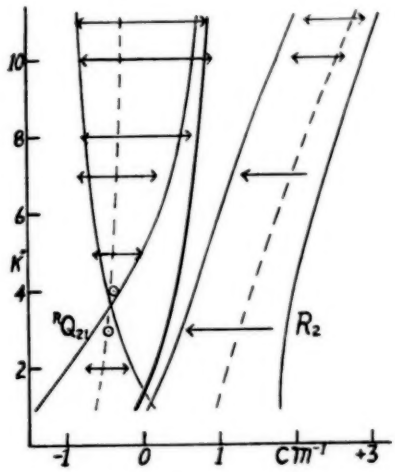


Fig. 4. Zeeman patterns of the ${}^{R}Q_{21}+R_2$ lines of the ZnH band for $H=16{,}400$ gauss. All displacements are with reference to the no-field ${}^{R}Q_{21}$ positions. Note particularly the observed constriction in the width of the ${}^{R}Q_{21}$ block of components at $K''=3$ and 4.

cluded jumps $\Delta J=\pm 2$, and hence with increasing interaction energy of the spin in the molecule (increasing importance of J), these intensity differences evidence the decreasing effectiveness of the field in breaking down this internal coupling.

Because of the relative peculiarities of the course of the magnetic energy levels in the ${}^2\Pi_{1/2}$ and ${}^2\Sigma$ states in Fig. 1 at the low K values, the computed pattern blocks for transitions such as ${}^{R}Q_{21}$ show a constriction at about $K''=3$ or 4. Fig. 4 illustrates this phenomenon for this particular branch. Experimentally it is found that the block of components is of about the same width for ${}^{R}Q_{21}(2\frac{1}{2})$ and $(5\frac{1}{2})$, and that both of these are distinctly broader than the ${}^{R}Q_{21}(3\frac{1}{2})$ and $(4\frac{1}{2})$ patterns. The latter are, as a matter of fact, no broader on the plate than the no-field comparison lines. Similar excellent agreement with the theory on this point is found in the ${}^{S}R_{21}$ patterns. The degree of fit between the predicted and observed patterns for the measurable ${}^{R}Q_{21}$ patterns can be judged from Fig. 4. It should be mentioned that measurement of the exact position of the edges of many of these broad patterns

is difficult, and that therefore exact coincidence with the predicted outlines of the patterns is not to be expected.

The Zeeman Effect in the CdH Bands

The magnetic energy levels for both of the CdH states involved in the λ4509 band may be computed from Eqs. (1), (2), and (3), with $B = 6.0\,\text{cm}^{-1}$, $\lambda = +169\,\text{cm}^{-1}$, and the ρ-type doubling in the $^2\Sigma$ state as given above. The results plotted as in Fig. 1 are similar to those for ZnH, with the following main points of difference. In the $^2\Pi_{\frac{1}{2}}$ state the two extreme levels start at $K = 2$ exactly as for ZnH, but rather than crossing each other at a low K value, they converge symmetrically and more gradually to about zero separation at $K = 17$. The two extreme levels in the $^2\Pi_{\frac{3}{2}}$ state do not diverge so rapidly, the interval of each from the no-field position being but $0.13\,\text{cm}^{-1}$ at $K = 15$. And finally in the $^2\Sigma$ state the two groups of magnetic levels diverge of course more rapidly because of the larger ρ-type doubling, while

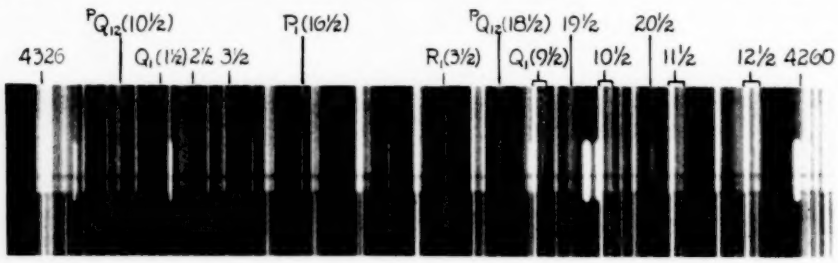


Fig. 5. Reproduction of the portion of the (0, 0) ZnH band between the $^2\Pi_{\frac{3}{2}}$ and $^2\Pi_{\frac{1}{2}}$ origins at $H = 16{,}400$ gauss. The doublet structure of the first Q_1 and R_1 lines is evident. The high frequency weaker component of the patterns $Q_1(9\frac{1}{2})$ to $(12\frac{1}{2})$ is opposite the $^QR_{12}$ satellite in each case. Notice the breadth of these satellite field blocks (also $^PQ_{12}(18\frac{1}{2})$, $(19\frac{1}{2})$, and $(20\frac{1}{2})$) as compared to that of the $P_1(16\frac{1}{2})$ line.

the two curving levels with $M = -(K - \frac{1}{2})$ swing around more sharply in the region of low K values and at $K = 7$ are already spaced at a distance $\Delta\nu_n$ from the no-field positions.

The main differences between the Zeeman patterns of the CdH and ZnH lines due to these variations in the level scheme just mentioned are: first, that for CdH the two blocks of patterns arising from the transitions to the two sets of levels in the $^2\Sigma$ state diverge more rapidly from each other as K increases; second, that all of the blocks of components except those associated with the $^RQ_{21}$, $^QP_{21}$, and $^SR_{21}$ transitions are for $K > 5$ about $2\Delta\nu_n$ in width, while the patterns of the lines of these three branches approach this width gradually as K approaches the value 17; third, the vanishingly weak intensity of some of the blocks of components already discussed in the ZnH section of this paper.

Comparison of the computed and observed patterns for one set of transitions, $R_2 + {}^RQ_{21}$, will suffice. This comparison is made in Table III and in Fig. 6 for the first seven lines in these branches. The calculated and observed widths and positions agree about as well as could be expected. Particularly

satisfying is the observation of the constriction in the width of the ${}^{R}Q_{21}$ block at $K''=3$. The ${}^{R}Q_{21}$ patterns are stronger than the corresponding R_2 patterns, even more so than the ratio of the intensities of the no-field comparison lines. This is due perhaps entirely to the fact that the ${}^{R}Q_{21}$ blocks are narrower. Also the low frequency edges of the ${}^{R}Q_{21}$ patterns and the high

TABLE III. *CdII (0,0) band. Zeeman patterns of $R_2+{}^{R}Q_{21}$ lines. $H=16{,}400$ gauss. Refer to Fig. 6. All displacements are with respect to R_2 no-field position. * indicates presence of overlapping pattern.*

K''	R_2 block		${}^{R}Q_{21}$ block	
	Obs.	Calc.	Obs.	Calc.
1	−2.04*to +1.52*	−2.10 to +1.50	same. see note below	
2	−0.88 to +1.23	−0.75 to +1.30	−2.00*to −1.31*	−2.05 to −1.35
3	−0.87 to +1.06	−0.80 to +1.15	−2.33 to −1.97	−2.36 to −1.80
4	−0.60 to +1.00	−0.82 to +1.05	−2.25*to −2.92*	−2.28 to −3.05
5	−0.70 to +0.91	−0.82 to +0.95	−2.68*to −3.58*	−2.80 to −3.75
6	−0.74 to +0.80	−0.81 to +0.92	−3.52 to −4.30	−3.30 to −4.37
7	−0.84 to +0.84	−0.80 to +0.90	−3.87 to −4.85	−3.87 to −5.00

Note: The measurements affected by overlapping patterns are made by noting the patterns of the lines of the same branch immediately preceding and following the line causing the fusion, and so determining the excess attributable to the R_2 or ${}^{R}Q_{21}$ line. These estimates are of course somewhat inaccurate.

frequency edges of the R_2 patterns are the stronger as they should be because of the slightly greater concentration of components on these sides of the blocks respectively. The ${}^{R}Q_{21}(12\frac{1}{2})$ pattern is of uniform intensity over all, however, for the $M'=M''=0$ line dividing equally the total number of components becomes centered for the higher K levels.

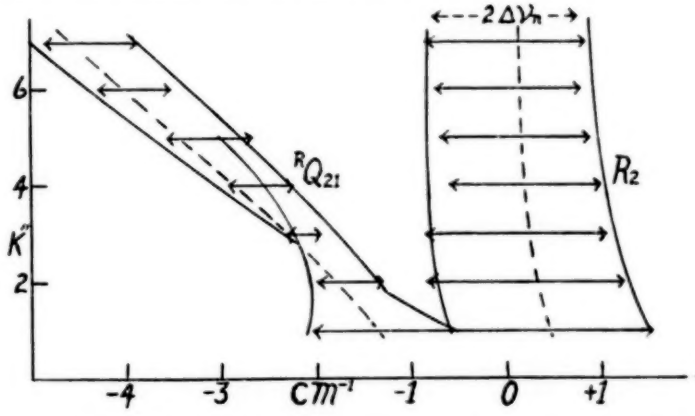


Fig. 6. Comparison of computed and observed Zeeman patterns for the $R_2+{}^{R}Q_{21}$ lines of the CdH (0, 0) band at $H=16{,}400$ gauss. Note the rapid assumption of the width $2\Delta\nu_n$ for the R_2 block of components as K increases, the narrowing of the ${}^{R}Q_{21}$ block at $K''=3$, and the rapid separation of the two blocks with K due to the larger ρ-type doubling.

Many more similar comparisons of these magnetic field effects in these two band systems with the predicted appearance of the patterns from Hill's quantum mechanics formulas might be presented. Enough of the details unique to these bands have been given, however, to demonstrate the good agreement between the theory and the observations.

THE PRINCIPLE OF IDENTITY AND THE EXCLUSION OF QUANTUM STATES

By Gilbert N. Lewis
Department of Chemistry, University of California

(Received August 11, 1930)

Abstract

In classical physics it was assumed that a system of two bodies, both having the same mechanical and electrical properties, would show the same behavior regardless of their complete identity. This assumption, that gave rise to the paradoxes of Gibbs and of Einstein which are here discussed, must be abandoned in atomic physics. We must distinguish sharply between identical particles and those that are only quasi-identical, in the sense that they are the same in mass, charge and the potential fields to which they give rise. All of the results which have hitherto been obtained in quantum mechanics are valid even if we assume that two particles such as two electrons or two helium nuclei are only quasi-identical. The postulate of complete identity is a new mathematical condition which, operating simultaneously with the wave equation and its boundary conditions, restricts the number of possible solutions of the equation and thus gives rise mathematically to a number of exclusion laws such as those which have been empirically obtained.

The assumption of identity is equivalent to the mathematical assumption that the ordinary configuration space is degenerate, in the sense that one portion is a mere reflection of another. By this method it is shown that if the wave function is a scalar, the laws of exclusion are obtained without ambiguity. Thus a perfectly symmetrical diatomic gas in its simplest electronic state can have only even rotational levels. It is pointed out that the experimental results with $O_2(O_{16}O_{16})$ and He_2 show that these atoms exist in a single *nuclear* state. It is predicted that with heavier atoms both even and odd states will be found owing to the existence of isobars. The monatomic gas is then discussed and it is shown that, if the Schrödinger function is a scalar, only the Bose-Einstein statistics are possible.

The phenomenon of spin introduces a vector character into the wave function. If we assume that the spin factor changes sign upon reflection when the number of units of spin is odd but not when the number of units of spin is even, we have all the cases of exclusion now known, including the case of the nitrogen molecule. As an example of the method, the problem of the helium atom is discussed.

THERE is at present a great interest in the various rules for the exclusion of certain quantum states, of which rules Pauli's was the prototype; and many speculations are rife concerning those quantum states which are supposed to issue from the mathematics of quantum mechanics, but of which it is said that they "do not exist in nature." I have therefore decided to present, before all their applications have been worked out, certain ideas through which I wish to show that all such exclusion rules become direct mathematical consequences of quantum mechanics when full consideration is given to the identity of certain elementary particles.

In the last century there was no convincing evidence that any two atoms of a given species could be assumed to be exactly alike, but we are now convinced, although the number of atomic species has been largely increased

by the discovery of isotopes, that the individuals within any one species are in all respects identical. This identity has been tacitly recognized without a complete realization of the remarkable consequences to which it leads. For the first intimation of these consequences we must go back to a notable discussion by Willard Gibbs.

The Paradoxes of Gibbs and of Einstein

In "The Equilibrium of Heterogeneous Substances," Gibbs, obtaining an expression for the increase of entropy when two gases are mixed, commented as follows: "It is noticeable that the value of this expression does not depend upon the kinds of gas which are concerned, ... except that the gases which are mixed must be of different kinds. If we should bring into contact two masses of the same kind of gas, they would also mix, but there would be no increase of entropy." He goes on to say, "Now we may without violence to the general laws of gases which are embodied in our equations suppose other gases to exist than such as actually do exist, and there does not appear to be any limit to the resemblance which there might be between two such kinds of gas. But the increase of entropy due to the mixing of given volumes of the gases at a given temperature and pressure would be independent of the degree of similarity or dissimilarity between them. We might also imagine the case of two gases which should be absolutely identical in all the properties (sensible and molecular) which come into play while they exist as gases either pure or mixed with each other, but which should differ in respect to the attractions between their atoms and the atoms of some other substances, and therefore in their tendency to combine with such substances. In the mixture of such gases by diffusion an increase of entropy would take place, although the process of mixture, dynamically considered, might be absolutely identical in its minutest details (even with respect to the precise path of each atom) with processes which might take place without any increase of entropy."

As the increase of entropy remains constant when the two gases that are mixed continuously approach complete identity, then by the doctrine of limits, the same increase must be found in the limiting case of identity; but since there is no change of entropy on mixing two identical gases, we are here confronted with a real paradox. In other words, some one link in the chain of reasoning must be fallacious. Despite the penetrating genius of Gibbs, it would have been too much to expect of him in his time to resolve this paradox, but since the advent of quantum theory it is easy to see where the difficulty lies. The false step was made in assuming that two species could approach identity by a continuous path. We are now aware that even atoms of the same species can not be regarded as identical if they are in different quantum states, since these quantum states constitute a discrete series.

While the formal paradox is thus resolved, we are left with a feeling that the phenomenon discussed by Gibbs is still antagonistic to some underlying physical principle; for quantum states, while always distinguishable from

one another, may lie very close together. We have long been taught that the behavior of a physical system is completely determined by such quantities as mass and potential, and we therefore feel that systems which are almost exactly alike with respect to such quantities must behave nearly the same. These ideas brought over from the older physics we must now abandon.

The paradox of Gibbs is important not only as the first intimation of a new kind of principle operating in physical science, but also because when it is stated in modern terms it suggests a general rule that is immediately applicable to problems of atomic and molecular structure. The entropy of a system under given conditions is a measure of the number of its quantum states that are compatible with these conditions. When we find, other things being equal, that the entropy of a system is greater when it is composed of two kinds of particles than when it is composed of one kind, this means that the number of quantum states is reduced in the case of identical particles This rule holds, for reasons that we shall see presently, not only for the great aggregates of particles with which thermodynamics deals, but also for the simplest atomic systems.

Let us next turn to a problem presented by Einstein, who, in his work on the monatomic gas, discovered the remarkable deviation from the laws of the perfect gas which must result when the same individual quantum numbers may be assigned to two or more particles. I quote from the closing paragraph of his paper.[1] "Finally I may call attention to a paradox, the solution of which I have sought without success. There is no difficulty in treating, by the method here given, the case of the mixing of two different gases. In this case each molecular species has its separate "cells." Thence follows the additivity of the entropies of the components of the mixture, and each component behaves, with respect to molecular energy, pressure, and statistical distribution, as if it alone were present. Therefore a mixture with the molecular numbers n_1, n_2, where the molecules of the first and second species differ indefinitely little from each other, (especially with respect to the molecular masses m_1, m_2) exhibits, at a given temperature, a different pressure and distribution from a single gas with n_1+n_2 molecules of practically the same mass, and at the same volume. This, however, appears as good as impossible."

Here is a paradox only if we take over from classical physics the idea that the behavior of a system is determined by such quantities as mass and potential alone. This, however, is the very idea that we have now agreed to abandon. The equations of Einstein, like the equations of Gibbs, must be accepted. It must be admitted that the equation of state for a mixture of isotopes is unlike that for either pure isotope, although at present this seems beyond the limits of experimental proof. This is true not only of two nearly identical substances, but as Dr. Mayer and I have shown,[2] it is also true for molecules of the same substances if they are in different quantum states. Thus a gas composed of normal and one composed of metastable

[1] Einstein, Berlin Ber. 267 (1924).

[2] Lewis and Mayer, Proc. Nat. Acad. **15**, 208 (1929).

mercury atoms may exhibit the same pressure under like conditions, but a mixture of the two must have a different pressure.

Identical and Quasi-Identical Systems

An electron in a sodium atom is distinguishable from an electron in a chlorine atom, but if we imagine the two interchanged, the resulting system differs from the original in no way whatsoever. In the same sense, two helium atoms may be regarded as intrinsically identical, but not always two lithium atoms, for one may have the mass 6, and the other the mass 7. We may even have two isotopes of nearly, or exactly, the same mass. Such "isobars" are found in the radioactive series where two atoms have the same atomic number and, except for their small differences in energy, the same atomic weight, and yet differ in origin and in later radioactive behavior. Two such particles may be regarded as composed of the same number of protons and electrons in a different arrangement, or, as we now say, they represent the same system in two different quantum states. We might even find two atoms with precisely the same mass (analogous to optical isomers) or, in other words, the lowest energy level of the nucleus may be degenerate. Nevertheless, two such atoms, or two such states of an atom, must not be regarded as identical, though they may be so regarded for most purposes.

When two non-identical particles have the same, or nearly the same, values for those quantities which enter into the quantitative statement of a certain problem, they may then be called quasi-identical. Whether those values could be precisely the same for non-identical particles is doubtful. Suppose, for example, that in a certain problem we are concerned with the mass and charge of two particles, and with the potential field to which each gives rise. In the case of two isobars of identical mass and charge, the potential field would presumably be different at very small distances from the atomic center. If we held those prejudices which made the equations of Gibbs and of Einstein seem paradoxical, we should expect to detect such lack of complete identity only by measurements of extreme refinement, calculated to show the minute differences to which we have referred. On the other hand, by using the principle which I am here developing, the roughest sort of observation, such as the mere inspection of a spectrogram, may suffice to decide as to the identity of the particles.

The Quantum States of a Diatomic Molecule

When the electronic states are not such as to produce complications, the rotation and vibration states of a diatomic molecule have been treated with complete satisfaction by the methods of Schrödinger as applied to a two-body problem. Let us consider two quasi-identical atoms, A_1 and A_2, with the same mass, m, and a potential field for each atom which depends only upon the distance from the other. The Schrödinger equation then reads

$$\frac{\partial^2\psi}{\partial x_1^2}+\frac{\partial^2\psi}{\partial y_1^2}+\frac{\partial^2\psi}{\partial z_1^2}+\frac{\partial^2\psi}{\partial x_2^2}+\frac{\partial^2\psi}{\partial y_2^2}+\frac{\partial^2\psi}{\partial z_2^2}+\frac{8\pi^2 m}{h^2}(E-V)\psi=0 \quad (1)$$

where the coordinates x_1, y_1, z_1, and x_2, y_2, z_2 are associated, respectively, with the first particle and with the second, and ψ is a (real or complex) scalar function of the coordinates.

Such an equation, even when the potential V is given as an explicit function of the coordinates, does not suffice to determine ψ, but only restricts the range of its possible values. This range is further restricted if we impose such further specifications as are known as boundary conditions. In the present case we employ the condition that ψ shall remain finite and continuous. The continuous range of the solutions of Eq. (1) is then reduced (when $E<0$) to a discontinuum of discrete proper-functions. Any additional condition which operates simultaneously with Eq. (1) and its boundary conditions will in general further diminish the range of the solutions. This lies at the very root of mathematical reasoning. In the present case, where the solutions are already reduced to a discrete series, the assumption of identity will in general eliminate some of these solutions.

We need not go through the familiar process of integrating Eq. (1). but merely recall that the coordinates are first transformed to coordinates of the center of gravity and the relative coordinates x, y, z, where $x=x_1-x_2$, etc. The equation may then be split into two equations, of which the second involves only x, y and z. Again, transforming to ordinary polar coordinates, r, ϕ and θ, the equation is further separated and integrated. The whole proper-function ψ now appears as the product of several partial functions of which those that depend upon ϕ alone and upon θ alone are

$$\Phi = e^{im\phi}, \tag{2}$$

$$\Theta = \frac{(2j)!\,\sin^m\theta}{2!(j-m)!}\left(\cos^{j-m}\theta - \frac{(j-m)(j-m-1)}{2(2j-1)}\cos^{j-m-2}\theta + \cdots\right). \tag{3}$$

where j and m are integers.

We must now find the consequences of imposing the *additional* mathematical condition that the two particles are not merely quasi-identical, but are identical. This can affect the result not by any modification of the solutions previously obtained but only by an elimination of some of them. In proceeding to the mathematical calculation a serious difficulty appears at the outset because in setting up Eq. (1) we associated one particle with the coordinates x_1, y_1, z_1, and the other with the coordinates x_2, y_2, z_2. As the two particles are indistinguishable this is no longer possible. It might be possible to devise a new mathematical method in which the identity of the particles is recognized from the beginning, but until such a method has been invented we must start as before with the fiction of quasi-identical particles and later make a correction for the consequences of this fiction. This must be done by recognizing that the six dimensional space $x_1 \ldots . z_2$, is a degenerate or "mirror" space, in the sense that the conditions in one-half of the space must be an exact reflection of those in the other half. Thus, any condition prevailing in the point $x_1=a$, $y_1=b$, $z_1=c$ and $x_2=d$, $y_2=e$, $z_2=f$ must be exactly reflected at the point $x_1=d$, $y_1=e$, $z_1=f$ and $x_2=a$, $y_2=b$ and $z_2=c$.

This may be illustrated in the simple two-dimensional space represented in Fig. 1. In this space of the coordinates x_1 and x_2 the diagonal dotted line acts as the mirror and all conditions on one side of this line are automatically reproduced by reflection on the other side. Thus the point P'' is the reflection of the point P'. A scalar function of position must be the same even to sign at the two points. On the other hand, this need not be true of a vector function of position. Suppose that we are considering a planar vector or 2-vector (which in this two dimensional space acts as a pseudo-scalar) of which the magnitude is represented by the area of a circle and its direction by an arrow, as in the figure. It is evident that the vector is reflected as a vector of the same magnitude but of opposite direction.[3]

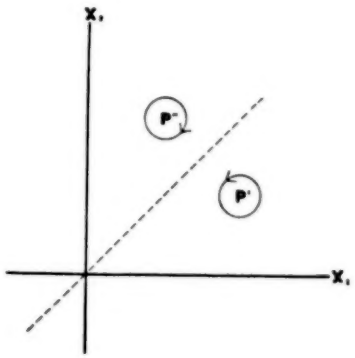


Fig. 1.

Returning to the diatomic molecule we see that the principle of identity requires that the scalar function ψ, must be the same, even to sign, at two mutually reflecting points. If ψ is broken into several factors, then, because of the variety and the arbitrariness of the ways in which a quantity may be factored, our rule of invariance to reflection need not apply to the individual factors, but only to their product.

When as before we transform to new coordinates all the degeneracy of the space remains in the space of the relative coordinates. $x=x_1-x_2$ can not be distintuished from $-x=x_2-x_1$, so that every point has its reflection at a diametrically opposite point. When, therefore, we finally change to polar coordinates, the value of ψ at a given point (ϕ, θ) must be the same as at the reflected point $(\phi'=\pi+\phi, \theta'=\pi-\theta)$. If now we examine Eq. (2) and (3) we see that Φ changes sign, if ϕ is replaced by $\pi+\phi$, whenever m is odd, and Θ changes sign, if θ is replaced by $\pi-\theta$ whenever $j-m$ is odd. Therefore, in order that ψ have the same sign at reflected points it is necessary that $j-m$ be even when m is even and that $j-m$ be odd when m is odd; thus, j must always be even. In other words, of all proper-functions which are

[3] While the pseudo-scalar in two dimensional space (which may be regarded as the vector product or outer product of two linear or 1-vectors) changes sign on reflection, in four dimensional space the pseudo-scalar does not change sign on reflection. In six dimensional space the pseudo-scalar again changes sign.

solutions of the former set of mathematical conditions only those with even values of j are solutions of the new set of mathematical conditions.

These deductions are completely corroborated by experiment. Not only from spectroscopic observations but also from the entropy determination of Giauque and Johnston,[4] we know that the molecule $O_{16}O_{16}$ has only half as many quantum states as a molecule of unlike particles. On the other hand $O_{16}O_{17}$ and $O_{16}O_{18}$, discovered by Giauque and Johnston, show all the states corresponding to the solution of the quasi-identical problem. Moreover, such particles as O_2 and He_2, with identical particles, have in their simplest electronic states only the *even* rotational states. This agreement of our deductions with experiment is to be regarded not as a corroboration of our mathematical method, which, I believe, requires no such corroboration, but rather as a justification of the application of the *scalar* equation, Eq. (1), to the simple problem of two bodies.

The experimental results go farther and give us an insight into the nucleus itself, for if the atoms O_{16} and He existed to an appreciable degree in more than one nuclear quantum state, the lines now missing in the molecular spectra would have been found, and their intensities would inform us concerning the relative amounts of the different isobars. Such isobars were not perhaps to be expected in these simple nuclei, but in elements of higher atomic weight, where many isotopes have been discovered, we may expect to find lines in the spectra of the diatomic molecules which give definite evidence of the existence of isobars.

It can not be too strongly emphasized that the conclusions hitherto obtained are in no way dependent upon the particular choice of the Schrödinger equation. They will be entirely similar if we use the relativistic form of that equation, or indeed, any equation whatsoever that defines a *scalar* function of position and that determines proper-functions, with the aid of its boundary conditions.

The Monatomic Gas

Let us consider next the simplest of all 2-body problems: namely, that of two quasi-identical particles in a box free from potential fields except at, or very near, the boundaries. The new boundary condition is that ψ shall be zero outside the box. If we assume an equation such as Eq. (1) in which V is omitted, it is easily separated by writing $\psi = \psi_1\psi_2$. We thus obtain two sets of partial proper-functions so that, if ψ_{1m} represents the first particle in the m'th quantum state, and ψ_{2n} represents the second in the n'th state, the product of the two is a proper-function for the whole system.

On account of the quasi-identical nature of the two particles $\psi_{1m}\psi_{2n}$ is associated with the same energy as $\psi_{2m}\psi_{1n}$ and therefore any linear aggregate of these two is also a solution of the equation. It is customary to take as the two orthogonal solutions

$$\psi^+ = \psi_{1m}\psi_{2n} + \psi_{2m}\psi_{1n};\ \psi^- = \psi_{1m}\psi_{2n} - \psi_{2m}\psi_{1n}. \tag{4}$$

[4] Giauque and Johnston, J. Am. Chem. Soc. 51, 3300 (1929).

Let us now once more replace the condition of quasi-identity by the condition of identity. There are no longer two degenerate states, $\psi_{1m}\psi_{2n}$ and $\psi_{2m}\psi_{1n}$; the single state is completely specified by stating that one particle is in state m and the other in state n. To proceed more formally, we know again that the space $x_1 \cdots z_2$ is a degenerate or "mirror" space in which every point P' has its reflected point P''. Owing to the identity of the particles $\psi_{1m}\psi_{2n}$ at P' is the same as $\psi_{2m}\psi_{1n}$ at P'', and $\psi_{2m}\psi_{1n}$ at P' is the same as $\psi_{1m}\psi_{2n}$ at P''. Of the two functions ψ^+ and ψ^- of Eq. (4), the former remains unchanged during reflection, the latter changes sign and must therefore be excluded if the whole wave function is a scalar.

This reduction of the two sets of proper-functions to the single function ψ^+ corresponds to the smaller entropy of a system of identical particles discussed by Gibbs, and constitutes the difference between the Bose-Einstein statistics and the statistics of Boltzmann.

If, however, we are dealing with a system in which ψ is not a scalar but changes sign on reflection (as in the case exemplified in connection with Figure 1) then of the functions ψ^+ and ψ^- the former would have been excluded and only the latter allowed. It is to be noted that, because of its form, ψ^- always disappears when both particles are in the same quantum state; that is, when $m = n$. We therefore have a system following the Fermi statistics.

It is easy to generalize to a large number, N, of identical particles. Here the configuration space becomes highly degenerate and may be divided into $N!$ equal pyramidal regions such that if conditions are given throughout one of these regions they are obtained by reflection in all the other regions. Instead of the pair of mutually reflecting points in the previous case we now have $N!$ mutually reflecting points. Furthermore, in the system of N particles the $N!$ degenerate states of the quasi-identical particles are reduced to a single state in the case of identity.

Any one of the quasi-identical states may be obtained from any other by a series of single reflections, each of which may be interpreted as the interchange of a single pair of particles. They may be divided into two groups according as they may be obtained from a chosen state by an even or odd number of reflections. The sum of one group added to the sum of the other group gives the single function that we have called ψ^+, and the sum of one group subtracted from the sum of the other gives ψ^-.

If the whole wave function is a scalar, ψ^+ is the only function allowed and the number of quantum states corresponds to the Bose-Einstein statistics. If, however, the whole wave function changes sign with each single reflection, ψ^- is the only function allowed. The terms in ψ^- vanish pair by pair whenever two of the particles are in the same state and we thus have the Fermi statistics. In neither case do we have any of the so called "mixed types" of proper-functions.

The Problem of Spin

The principle of identity tells us without ambiguity how many of the various sets of quantum states obtained in the case of quasi-identity are ex

cluded in the case of identity. If the nature of the wave function is known it tells further just which sets of quantum states are excluded. Thus if the wave function is a scalar we have seen how simply we are led to the exact exclusion laws.

Further than this our principle can not go alone, but it will enable us to codify in a simple manner certain laws derived from experiment. The discovery of a fourth quantum number, the quantum number of spin, is incompatible with any equation of the type of Schrödinger's. It introduces a new and, in spite of the brilliant work of Dirac, a still mysterious element into quantum mechanics.

We shall assume (1) that in the absence of spin any system may be adequately interpreted by means of a scalar equation of the type of Schrödinger's. The principle of identity then states that the whole wave function remains unchanged upon reflection. Thus, for example, any particles, neutral or charged, that are not associated with spin obey the Bose-Einstein statistics.

We shall further assume (2) that every system may be described by means of a *scalar* wave function Ψ which, at least in certain limiting cases, may be regarded as the product of a function ψ, which satisfies an equation of the Schrödinger type, and a quantity σ, which may be called the factor of spin, and may be regarded as the product of the spin factors of the individual particles.

Our final assumption (3), which now generalizes assumption (1), is that for particles with an even number of units of spin σ is unchanged by reflection, but that for particles having an odd number of units of spin any product of the individual spin factors changes sign with each single reflection (interchange of two identical particles). As far as I have at present been able to see, these simple assumptions cover all known cases of the exclusion of quantum states, including the case of the nitrogen molecule. In thinking over the consequences of these assumptions, I have had the benefit of many discussions with Professor G. E. Gibson, to whom I wish to express my gratitude.

I shall give only one example of the application of these rules to atomic problems. If, in considering the helium atom, we should make the provisional assumption that the two electrons are only quasi-identical, we would obtain all of the results obtained by Heisenberg in his treatment of the problem, but when we introduce the further condition of identity one set of quantum states is excluded. Let us see what these quantum states are.

With Eq. (1), the center of gravity of the atom being taken as the origin and the possibility of motion of the nucleus being neglected, we may neglect that part of V which depends upon the interaction of the two electrons, since we are now concerned not with the energies of the quantum state, but only with their number and characterization. The equation can then be separated into two similar "hydrogen-like" parts and each leads to its own set of proper-functions so that any one from one set multiplied by any one from the other gives a solution of Eq. 1. Proceeding as in a previous case, we obtain the double set of functions given in Eq. 4, ψ^+ and ψ^-.

Since, however, each electron has one unit of spin the whole wave function of the system is not ψ but $\Psi=\psi\sigma$, where σ is the product of the two individual spin factors, either of which may have one of the two values α or β. Thus, treating σ in the ordinary way as a proper-function, it has four possible values which may be written as (a) $\alpha_1\alpha_2$; (b) $\beta_1\beta_2$; (c) $\sigma^+=\alpha_1\beta_2+\alpha_2\beta_1$; (d) $\sigma^-=\alpha_1\beta_2-\alpha_2\beta_1$.

So far we have been assuming provisionally that the two electrons are only quasi-identical. We now introduce the condition of identity and employ the rule that each product of the individual spin factors changes sign upon reflection, so that when we pass from a point P' to its reflected point P'', $\alpha_1\alpha_2$ goes into $-\alpha_2\alpha_1$, $\beta_1\beta_2$ into $-\beta_2\beta_1$, $\alpha_1\beta_2$ into $-\alpha_2\beta_1$ and $\alpha_2\beta_1$ into $-\alpha_1\beta_2$. Of the four values of σ given above the only one that does not change sign upon reflection is σ^-.

Now the whole wave function Ψ, must not change sign; therefore, σ^- can be combined only with ψ^+ and the other three values of σ can be combined only with ψ^-. We thence obtain, by methods too familiar to repeat, the quantum states of the helium atom which are experimentally found; and our rule of exclusion covers in this, as it will in other cases, the Pauli principle.

ON THE WENTZEL-BRILLOUIN-KRAMERS APPROXIMATE SOLUTION OF THE WAVE EQUATION

By L. A. Young and G. E. Uhlenbeck

Department of Physics, University of Michigan

(Received August 25, 1930)

Abstract

A discussion of the Wentzel-Brillouin-Kramers method of obtaining an approximate solution of the wave equation is given from the point of view that it forms a link between the quantum theory of Bohr and the new quantum mechanics. This becomes clear when one compares the probability distributions (see Figs. 2 and 3) and calculates the mean values $\overline{r^k}$ (see Table I). It is shown that for high quantum numbers these become the classical values. The method leads also, as shown by Zwaan, to quantization by the classical phase integrals with the use of half-integer quantum numbers. In the last section the method is applied to the charged shell atom model. It is shown that the condition of smooth joining of the wave function is practically equivalent to half-integer quantization of the sum of the inner and outer phase integrals. Of course there is no longer a sharp distinction between penetrating and non-penetrating orbits. The Landé formula for the doublet separations is derived. The value of $\psi^2(0)$, which occurs in the calculation of the hyperfine separations in ^2S states, is also given.

I. General Character of Wentzel-Brillouin-Kramers Approximate Solution

THE method proposed by Wentzel and Brillouin and later developed by Kramers and Zwaan[1] for obtaining an approximate solution of the one-dimensional wave equation[2]

$$\psi''(x) + Q^2(x)\psi = 0 \tag{1}$$

where $Q^2(x) = (8\pi^2 m/h^2)(E - V(x))$ in which E is the total energy and $V(x)$ the potential energy of the system consists, as is well known, in substituting into Eq. (1) an expression of the form

$$\psi = \exp\left\{\frac{2\pi i}{h}\left(S_0 + \frac{h}{2\pi}S_1 + \frac{h^2}{4\pi^2}S_2 + \cdots\right)\right\}$$

where S_0, S_1, S_2, $\cdots$ are functions of x to be determined. One finds easily by successively equating the coefficients of equal powers of $h/2\pi$ to zero that

$$S_0 = \pm \int^x Q dx$$

[1] G. Wentzel, Zeits. f. Physik **38**, 518 (1926); L. Brillouin, C. R. Juli, 1926; H. A. Kramers, Zeits. f. Physik **39**, 828 (1926); H. A. Kramers and G. P. Ittmann, Zeits. f. Physik **58**, 217 (1929); A. Zwaan, Utrecht Dissertation, 1929.

[2] It should be noticed that $Q(x)$ is proportional to the linear momentum $p(x)$ and therefore to the velocity $v(x)$.

$$S_1 = + \frac{i}{2} \log Q$$

One obtains therefore as a first approximation to a solution of Eq. (1)

$$\psi_{\text{appr}} = Q^{-1/2}\left\{c_1 \exp\left(\frac{2\pi i}{h}\int^x Q dx\right) + c_2 \exp\left(-\frac{2\pi i}{h}\int^x Q dx\right)\right\}$$

The integration constants c_1 and c_2 must now be so chosen that in regions where a motion was possible classically ($Q^2(x) > 0$) ψ has an oscillatory character while in regions where $Q^2(x) < 0$ we have solutions behaving as a linear combination of increasing and decreasing exponentials.

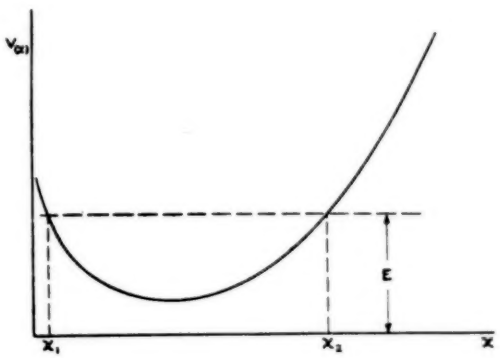


Fig. 1.

If, for example, we are considering a potential function of the type sketched in Fig. 1 the approximate solution will have the forms

$$\begin{aligned}
\psi_{appr} &= AQ^{-1/2} \cos\left\{\int_{x_1}^x Q dx + \delta\right\} \qquad x_1 < x < x_2 \\
&= BQ^{-1/2} \exp\left\{\int_x^{x_1} Q dx\right\} \qquad x < x_1 \\
&= CQ^{-1/2} \exp\left\{\int_{x_2}^x Q dx\right\} \qquad x > x_2.
\end{aligned} \tag{2}$$

These solutions have been so chosen that they become zero for $x = \pm\infty$ as they must if they are to correspond to an eigenfunction of the discrete spectrum of the wave equation.

When the expressions (2) are approximations to one and the same exact solution it is clear that there must be relations connecting the constants of amplitude and phase A, B, C, and δ. These relations are given by Kramers and Zwaan. They say that an increasing exponential solution like $\psi_{\text{appr}} = Q^{-\frac{1}{2}}\exp\{\int^x Q dx\}$ is "connected" with an oscillatory solution $\psi_{\text{appr}} = Q^{-\frac{1}{2}}$

$\cos\{\int^x Qdx-\pi/4\}$ while a decreasing exponential solution $\frac{1}{2}Q^{-\frac{1}{2}}\exp\{-\int^x Qdx\}$ must connect with an oscillatory function $Q^{-\frac{1}{2}}\cos\{\int^x Qdx+\pi/4\}$. These formulas completely determine δ but the constants A, B, and C are determined only within a common multiplicative constant which may be fixed by normalization.

The value of this approximate solution is two-fold; first, of course, it can lead to valuable results in cases where the exact solution is unknown or difficult to handle; in the second place it has the advantage that it is a link between the classical quantum theory of Bohr and the exact quantum mechanics. One may say that it shows how from the viewpoint of the latter, at least in the case of one-dimensional problems, the classical physicist ought to have treated his problems to obtain the best results. We will illustrate this in general:

a. by showing that the approximate solution leads to a determination of the energies by quantization of the classical action integrals with the sole difference that half-integer quantum numbers must be used instead of whole numbers.

b. by considering the probability distribution $P_{\text{appr}}=\text{const}\psi^2_{\text{appr}}$ and comparing it with the classical probability which is inversely proportional to the velocity, that is $P_{\text{class}}=\text{const}/v=\text{const}/Q$ as well as with the exact probability functions $P_{\text{exact}}=\text{const}\,\psi\bar{\psi}$.

In sections II and III for the case of hydrogenic atoms and the charged shell atom model we will show in more detail how the Wentzel-Brillouin-Kramers approximate solution fills the gap between the older and newer theories.

2. If we consider a potential function of the type sketched (Fig. 1) corresponding to an allowed energy E such that the classical motion was periodic and bounded by the points x_1 and x_2 we may apply the connection formulas at these points obtaining as our approximate solution for this region either

$$Q^{-1/2}\cos\left\{\int_{x_1}^{x} Qdx-\pi/4\right\}$$

or

$$Q^{-1/2}\cos\left\{\int_{x}^{x_2} Qdx-\pi/4\right\}$$

We shall require that these two expressions join smoothly at every point x of the interval x_1 to x_2 (or, in other words the two must be identical). Since the solutions are determined only within an arbitrary multiplicative constant we must have $\psi_1'/\psi_1=\psi_2'/\psi_2$ where ψ_1 and ψ_2 denote the two alternative cosine-like solutions for this region. This condition requires that

$$\tan\left\{\int_{x_1}^{x} Qdx-\pi/4\right\}=-\tan\left\{\int_{x}^{x_2} Qdx-\pi/4\right\} \tag{3}$$

This has the solution $\int_{x_1}^{x} Q dx = -\int_{x}^{x_2} Q dx + (n_x + \frac{1}{2})\pi$ from which we see

$$\int_{x_1}^{x_2} Q dx = (n_x + \tfrac{1}{2})\pi \tag{4}$$

This relation determines the allowed energies and from our previous remark concerning the relationship between $Q(x)$ and the linear momentum $p(x)$ we see that this is equivalent to quantization by the action integrals of the Bohr theory with the single difference that half-integers now appear instead of the integers of the older theory.

$$I = \oint p dq = (n + \tfrac{1}{2})h \text{ instead of } I = nh.$$

This was first made clear by Kramers and later expressed in practically the above form by Zwaan in his Utrecht dissertation. It is felt that this affords a justification for the half-integer quantization which was found empirically to be so successful in certain problems, particularly in the interpretation of the vibrational spectra of diatomic molecules.

3. In discussing the character of the approximate solution we shall speak of the derived probability functions rather than the solutions themselves. As unnormalized probability functions we have

$$\begin{aligned} P &= \frac{1}{4} Q^{-1} \exp\left\{ -2 \int_{x_1}^{x} Q dx \right\} \qquad x < x_1 \\ &= Q^{-1} \cos^2 \left\{ \int_{x_1}^{x} Q dx - \pi/4 \right\} \qquad x_1 < x < x_2 \\ &= \frac{1}{4} Q^{-1} \exp\left\{ -2 \int_{x_2}^{x} Q dx \right\} \qquad x > x_2. \end{aligned} \tag{5}$$

First of all we notice that in the region $x_1 < x < x_2$ the probability is a positive oscillatory function with an amplitude equal to Q^{-1}—exactly the classical probability function. The close relation of this probability to the corresponding classical quantity is brought out if we write

$$P = Q^{-1} \cos^2 \left\{ \int_{x_1}^{x} Q dx - \pi/4 \right\}$$

in the equivalent form

$$P = \frac{1}{2} Q^{-1} - \frac{1}{2} Q^{-1} \sin \left\{ 2 \int_{x_1}^{x} Q dx \right\}. \tag{6}$$

The first term in this expression is just the classical probability. In calculating mean values the second term will contribute little due to its alternating sign and we will obtain to a first approximation the classical result with half-integer quantum numbers.[3] In the limit of large quantum numbers the ratio of the contribution of the second term to that of the first will vanish.

[3] It is now clear that the classical analogue of the matrix element X_{nm} connected with the transition between two states characterized by, say, the energies E_n and E_m is given by

The probability functions in the tail regions ($x<x_1$, $x>x_2$) had no meaning in the Bohr theory and it may be shown that their contribution may be considered as a small correction to the results obtained from the region in which a motion was classically possible. In the exact wave mechanical theory, however, the tail regions are not so sharply marked off from the interval[4] x_1 to x_2 and it may be shown that the results obtained by using the approximate solution lie between those of the classical theory and those of the Schrödinger theory.

Summarizing we see that the approximate probability function $P_{\text{appr}}=c\psi^2_{\text{appr}}$ has in common with the exact probability $P_{\text{exact}}=c\psi\bar{\psi}_{\text{exact}}$ an oscillatory character in the region $x_1<x<x_2$ and the existence of exponentially decreasing tails. With the classical probability function it has in common that at x_1 and x_2 it becomes infinite[5] and that the amplitude of the oscillatory part is equal to the classical probability function. See Figs. 2 and 3.

II. Application to Hydrogenic Atoms

4. The three dimensional motion of a charged particle (electron) in a central Coulomb field is described by a wave equation of the form

$$\Delta_3\psi+\frac{8\pi^2m}{h^2}\left(E-\frac{Ze}{r}\right)\psi=0$$

where Δ_3 is the Laplace operator for three dimensions. Introduction of spherical polar coordinates allows the separation of this partial differential equation and the solution may be written $\psi=R(r)y_l(\theta,\phi)$ and y_l is a spherical harmonic of the l^{th} order. By the substitution $w=rR$ the radial equation is reduced to the form

$$w''+Q^2(r)w=0$$

where

$$Q^2(r)=\frac{2Z}{r}-\epsilon-\frac{l(l+1)}{r^2}.$$

In this expression Z/r is the potential energy of an electron at a distance r from a Z-fold charged positive nucleus, ϵ is the total energy of the electron, and l has the physical meaning of total angular momentum.[6]

$$X_{nm}=\int\frac{X}{(Q_nQ_m)^{1/2}}dx\Big/\Big[\int\frac{dx}{Q_n}\int\frac{dx}{Q_m}\Big]^{1/2}.$$

[4] From Eq. (1) we see that at the points x_1 and x_2 $\psi''(x)=0$ so that these points are the outer inflection points of the exact solution.

[5] ψ^2_{appr} is, however, still integrable.

[6] All quantities are measured in the rational units first proposed by Hartree (D. R. Hartree, Proc. Camb. Phil. Soc. **24,** 89, 1928). In this system of units energy is measured in multiples of the negative energy of the hydrogen atom in its lowest stationary state; distances in terms of a unit equal to the radius of the first Bohr orbit of hydrogen; and time in units $1/4\pi cR$ where c is the velocity of propagation of light in vacuum and R is the Rydberg constant in sec^{-1}.

The above equation is of the form considered in the preceding section and the approximate solution considered there will be applied to the determination of the properties of hydrogenic atoms. Naturally, we ask first as to whether or not the correct energies are obtained when half-integer quantization is employed. We must evaluate the integral

$$\int_{r_1}^{r_2}\left[\frac{2Z}{r}-\epsilon-\frac{l(l+1)}{r^2}\right]^{\frac{1}{2}}dr$$

and equate it to $(n_r+\frac{1}{2})\pi$. The above integral was well known in the older theory giving

$$\pi(Z/\epsilon^{1/2}-[l(l+1)]^{1/2})=(n_r+1/2)\pi$$

or

$$\epsilon=\frac{Z^2}{(n_r+\frac{1}{2}+[l(l+1)]^{1/2})^2}.$$

It is apparent that we do not obtain the familiar Balmer formula except in the limit of large l values. If, however we replace $l(l+1)$ by $(l+\frac{1}{2})^2$ the correct result is obtained.

We shall now investigate the properties of the approximate eigenfunctions more closely. In the region $r<r_1$ we may choose either the positive or the negative exponential. This corresponds to our freedom of choice of either of the two roots of the indicial equation of the exact solution. We consider only the root which gives a solution vanishing at the origin and which behaves for very small values of r like r^{l+1}. The approximate solution based upon the negative exponential behaves for r very small like $r^{[l(l+1)]^{\frac{1}{2}}+\frac{1}{2}}$. Again the replacement of $l(l+1)$ by $(l+\frac{1}{2})^2$ leads to an agreement between the approximate and exact solutions. This is considered to be a sufficient argument for the consistent use of half-integers not only for the radial quantum number but for the azimuthal quantum number as well. Hereafter we shall consider $Q^2(r)=2Z/r-\epsilon-(l+\frac{1}{2})^2/r^2$ instead of the expression previously used.

5. From the wave mechanics it has become clear that most if not all observable properties of an atomic system can be expressed in terms of mean values of certain functions of the coordinates of the system. In the case of hydrogenic atoms for example the mean values $\overline{r^k}$ where k is a positive or negative integer, are very important. If P represents the probability function as a function of r we may write

$$\overline{r^k}=\frac{\int r^k P dr}{\int P dr}.$$

This expression may be used for the calculation of mean values classically if $P=P_{\text{class}}$ and the range of integration is over that region of the variable

where $Q^2(r)>0$; wave mechanically if $P=P_{\text{exact}}$ and the range of integration is extended from 0 to ∞; and by the Wentzel-Brillouin-Kramers approximation method if the approximate probability functions (5) are used and the integration taken over the entire range of the variable. For the latter case we may write

$$\overline{r^k}=\frac{\int_0^{r_1}P_{\text{app}}^{(1)}r^k dr+\int_{r_1}^{r_2}P_{\text{app}}^{(2)}r^k dr+\int_{r_2}^{\infty}P_{\text{app}}^{(3)}r^k dr}{\int_0^{r_1}P_{\text{app}}^{(1)}dr+\int_{r_1}^{r_2}P_{\text{app}}^{(2)}dr+\int_{r_2}^{\infty}P_{\text{app}}^{(3)}dr}=\frac{I_k^{(1)}+I_k^{(2)}+I_k^{(3)}}{I_0^{(1)}+I_0^{(2)}+I_0^{(3)}}.$$

By referring to Eqs. (6) we see that

$$I_k^{(2)}=\int_{r_1}^{r_2}P_{\text{app}}^{(2)}r^k dr=\int_{r_1}^{r_2}\left(\frac{1}{2}P_{\text{class}}-\frac{1}{2}Q^{-1}\sin\left\{2\int_{r_1}^{r}Q dr\right\}\right)r^k dr=C_k-U_k$$

If we now consider

$$\overline{(r^k)}_{\text{class}}=C_k/C_0$$

we obtain

$$\overline{r^k}\cong\overline{r^k}_{\text{class}}\left(1+\frac{I_k^{(1)}-U_k+I_k^{(3)}}{C_k}-\frac{I_0^{(1)}-U_0+I_0^{(3)}}{C_0}\right). \tag{7}$$

The above integrals were evaluated after making certain simplifying assumptions. It can be shown that for high total quantum numbers the correction terms vanish. In general, using (7) we get results lying between the classical values with half integer quantum numbers and the exact values. As an example we have tabulated $\overline{r^2}$.

TABLE I. *Hydrogenic atoms.*

	Class	Class (half-integers)	W.B.K.	Exact
$n=1, l=0$	1.0	2.1	2.5	3.0
$n=2, l=0$	34.0	38.5	40.8	42.0
$n=2, l=1$	16.0	26.5	28.6	30.0

In Figs. 2 and 3 the radial probability functions for the 3*s* and 3*d* states of hydrogen respectively have been plotted as a function of r. These examples illustrate the relation of the Wentzel-Brillouin-Kramers approximation to the results obtained by the exact wave mechanics.

6. In treating the frequency splitting in hyperfine structure due to the interaction of the extra-nuclear electrons with the nuclear magnetic moment it becomes necessary to calculate the magnetic field at the nucleus due to the orbital motion of the electron. This may be shown, in a classical way, to be proportional to $\overline{r^{-3}}$ and it would seem probable that the wave mechanical treatment should follow the classical treatment closely. However from the

fact that the exact solution of the Schrödinger equation behaves in the neighborhood of the origin like r^{l+1} and the probability distribution like $r^{2(l+1)}$ it is clear that $\overline{r^{-3}}$ diverges for s-states ($l=0$). This difficulty is only cleared

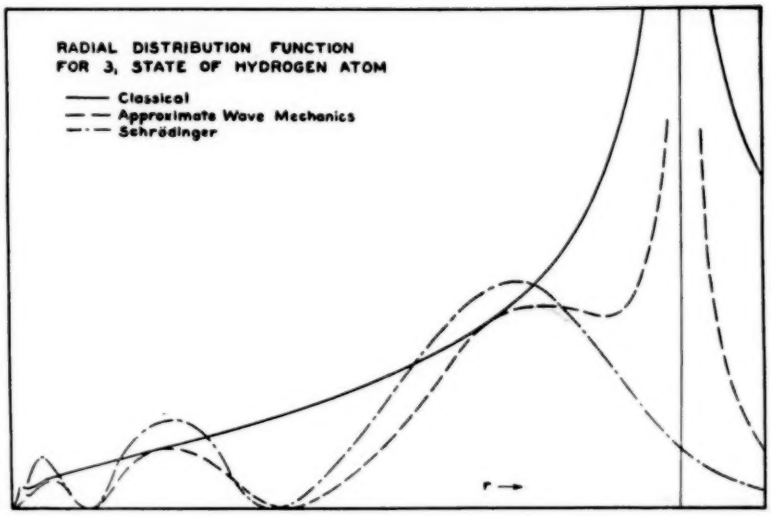


Fig. 2.

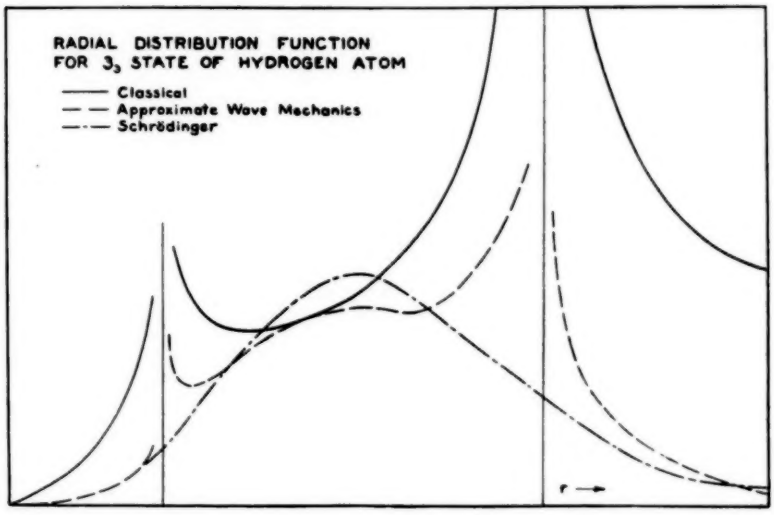


Fig. 3.

up when we treat the problem by means of the Dirac theory of the spin electron. Fermi and Hartree[7] have shown that for s-states we have only to replace $l\times\overline{r^{-3}}$ by $\frac{1}{2}R^2(0)$ where $R(0)$ is the value of the normalized radial eigenfunction at the origin, to obtain the correct results. Thus for $l\neq 0$ we have $\overline{r^{-3}}$:

[7] E. Fermi, Zeits. f. Physik **60**, 320 (1930); D. R. Hartree, Proc. Camb. Phil. Soc. **25**, 225 (1929).

Classical	Classical (half integers)	W. B. K.	Exact
$\frac{Z^3}{n^3(l+1)^3}$	$\frac{Z^3}{n^3(l+\frac{1}{2})^3}$	$\frac{Z^3}{n^3(l+\frac{1}{2})^3}(1+\text{corr.})$	$\frac{Z^3}{n^3 l(l+\frac{1}{2})(l+1)}$

While for $l=0$ we have:

"k" $\overline{r^{-3}}$		$\frac{1}{2}R^2(0)$	
Classical	Classical (half integers)	W. B. K.	Exact
$\frac{Z^3}{n^3}$	$\frac{4Z^3}{n^3}$	const. $\frac{Z^3}{n^3}(1+\text{corr.})$	$\frac{2Z^3}{n^3}$

For the case of $\frac{1}{2}R^2(0)$ by the Wentzel-Brillouin-Kramers approximation we are particularly interested to find out whether or not we obtain good agreement with the exact value since analogous methods when applied to the observed hyperfine structure of the alkali atoms can give valuable information concerning the structure of these nuclei.

For this case we may write

$$\frac{1}{2}R^2(0) = \frac{1}{2}\,\frac{\lim_{r=0}\frac{1}{4r^2Q}\exp\left\{-2\int_{r_1}^{r}Q\,dr\right\}}{I_0^{(1)}+C_0-U_0+I_0^{(3)}}.$$

Calculation of the numerator and using the values of the normalization integrals given in the appendix we find

$$\frac{1}{2}R^2(0) \cong \frac{\exp 2}{\pi}\cdot\frac{Z^3}{n^3}\left(1-\frac{1}{24n^2}\right)\left(1-\frac{I_0^{(1)}-U_0+I_0^{(3)}}{C_0}\right). \tag{8}$$

A numerical calculation for the states having $n=1,\ 2,\ \infty$ was made and the coefficients of Z^3/n^3 were found to be 2.60, 2.40, and 2.35. These are to be compared with the exact coefficient which is just 2.

III. Application to the Charged Shell Atom Model

7. The difficulties encountered in representing the properties of an alkali atom in the quantum mechanics are much the same as in the older theories. The calculation of the interaction between all the electrons of the atom seems to be a problem of the future. At present the only possibility seems to be to represent by some simple model as many of the properties of the atom as possible. The simplest mode of procedure is to assume that the core electrons act as a spherically symmetrical charge distribution about the nucleus and that the optical electron moves in the field of the nucleus and this charge distribution. Some progress has been made in the determination of the appropriate charge distribution by various investigators. The unfortunate thing about their results, however, is that the mathematical methods to which they lead are of a purely numerical or graphical type. Clearly one must idealize the problem if one is to obtain results of any generality. The simplest idealization that one can make is to consider the core electrons con-

centrated in one or more spherical surface distributions of charge about the nucleus. If we adopt this charged shell model our mode of procedure is quite clear.

Consider an atom with a Z-fold charged nucleus surrounded by a spherically charged shell at a distance r_0 from the nucleus. Suppose the total charge in this shell to be $Z-z$ electrons. We will then have to deal with two functions Q^2 instead of one; namely

$$Q_i^2(r) = \frac{2Z}{r} - \frac{2(Z-z)}{r_0} - \epsilon - \frac{(l+\frac{1}{2})^2}{r^2}$$

$$Q_0^{(2)}(r) = \frac{2Z}{r} - \epsilon - \frac{(l+\frac{1}{2})^2}{r^2}.$$

We shall seek solutions of the two differential equations

$$w_i'' + Q_i^2(r)w_i = 0 \qquad r < r_0$$

$$w_0'' + Q_0^2(r)w_0 = 0 \qquad r > r_0$$

which join smoothly at $r=r_0$ (a smooth join means $w_i'/w_i = w_0'/w_0$ at $r=r_0$). Freedom of choice of r_0 still remains and we shall suppose that it has been chosen to lie within the region where a motion was possible classically.

The first question we must consider is concerning half-integer quantization. To answer this consider the two approximate solutions

$$w_i = Q_i^{-1/2} \cos\left\{\int_{r_1}^{r} Q_i dr - \pi/4\right\}$$

$$w_0 = Q_0^{-1/2} \cos\left\{\int_{r}^{r_2} Q_0 dr - \pi/4\right\}$$

The smooth joining condition given above leads to

$$-\tan\left\{\int_{r_1}^{r_0} Q_i dr - \pi/4\right\} - \tfrac{1}{2}Q_i^{-2}(r_0)Q_i'(r_0)$$

$$= \tan\left\{\int_{r_1}^{r_2} Q_0 dr - \pi/4\right\} - \tfrac{1}{2}Q_0^{-2}(r_0)Q_0'(r_0)$$

(note that $Q_i(r_0)=Q_0(r_0)$) which by the introduction of the notation

$$\int_{r_1}^{r_0} Q_i dr - \pi/4 = \alpha$$

$$\int_{r_0}^{r_2} Q_0 dr - \pi/4 = \beta$$

$$\tfrac{1}{2}Q_0^{-2}(r_0)\{Q_0'(r_0) - Q_i'(r_0)\} = \eta$$

may be written

$$\tan\alpha + \tan\beta = \eta.$$

If $Q_0'(r_0) - Q_i'(r_0)$ were equal to zero (the inequality existing corresponds to the fact that the electric force has a discontinuity at r_0) we would have $\eta = 0$ whence

$$\tan \alpha + \tan \beta = 0$$

$$\alpha + \beta = \eta_r \pi$$

or

$$\int_{r_1}^{r_0} Q_i dr + \int_{r_0}^{r_2} Q_0 dr = (\eta_r + \tfrac{1}{2})\pi .$$

That is to the order of accuracy with which we may replace η by zero half-integer quantization is retained. The method used here should be compared with the corresponding procedure used in the treatment of hydrogenic atoms.

When η may not be considered zero it is necessary to consider

$$\tan \alpha + \tan \beta = \eta .$$

In case η is small we may write

$$\int_{r_1}^{r_0} Q_i dr + \int_{r_0}^{r_2} Q_0 dr = (\eta_r + \tfrac{1}{2})\pi + \tan^{-1} \eta .$$

The evaluation of this expression gives the possible energies for the charged shell model as a function of the quantum numbers and of r_0—that is a Rydberg formula which is the analogue of the Balmer formula for the terms of hydrogen-like atoms.

It is perhaps worth while to remark that the deviation from half-integer quantization found here is a defect of the charged shell model and that if we replaced it by a model using the charge distributions of Hartree or Fermi[8] we would again have half-integer quantization.

8. We can now use the approximate wave functions to calculate the quantities $\overline{r^k}$ for alkali-like atoms. All correction integrals will be neglected (in other words, the calculation is purely classical) since the value of r_0 is so uncertain, that greater accuracy is unwarranted. We have, then, to consider the integrals

$$\int_{r_1}^{r_0} \overline{r^k} Q_i^{-1} dr \quad \text{and} \quad \int_{r_0}^{r_2} r^k Q_0^{-1} dr$$

These may be replaced approximately by

$$\int_{r_1}^{r_1'} r^k Q_i^{-1} dr \quad \text{and} \quad \int_{r_2}^{r_2} r^k Q_0^{-1} dr$$

where r_2' and r_1' are the inner and outer roots of Q_0^2 and Q_i^2 respectively.

[8] D. R. Hartree, Proc. Camb. Phil. Soc. **24**, 89 (1928); E. Fermi, Zeits. f. Phys. **48**, 73 (1928).

Using the notation introduced in section II. for hydrogenic atoms we may write immediately:

$$\overline{r^k} = \frac{E^{-1/2}C_k^{(i)} + \epsilon^{-1/2}C_k^0}{E^{-1/2}C_0^{(i)} + \epsilon^{-1/2}C_0^{(0)}} \tag{9}$$

where we have put

$$E = \epsilon + \frac{2(Z - z)}{r_0}.$$

Application of this formula for $\overline{r^{-3}}$ gives

$$\overline{r^{-3}} = \frac{Zz^2}{n_{\text{eff}}^3(l + \frac{1}{2})^3} \tag{10}$$

where $\epsilon = Z^2/n^2_{\text{eff}}$ and we neglect $\epsilon^{-\frac{1}{2}}C_3^{(0)}$ against $E^{-\frac{1}{2}}C_3^{(i)}$ and $E^{-\frac{1}{2}}C_0^{(i)}$ against $\epsilon^{-\frac{1}{2}}C_0^{(0)}$. This gives the well-known Landé formula[9] for the doublet separations in optical spectra.

Following Schrödinger[10] we could choose r_0 as the perihelion distance of the parabolic orbits of hydrogenic atoms $r_0 = (l+\frac{1}{2})^2/2z$. It is felt, however that this choice of r_0 has lost some of its physical meaning since the development of wave mechanics.

A procedure analogous to that used for hydrogenic atoms for the determination of $\frac{1}{2}R^2(0)$ gives

$$\frac{1}{2}R^2(0) \cong 2.35 \cdot \frac{Zz^2}{n_{\text{eff}}^3}\left(1 - \frac{2z}{3Z} + \cdots\right).$$

Replacement of "$2Z^3/n^3$" by the above expression in the formula for the frequency separation in hyperfine structure of the ^2S states of alkali atoms[11] gives for the product of the Landé g-value of the nucleus and the mechanical moment of nucleus and electron

$$\begin{aligned} g(i)(i+\tfrac{1}{2}) &= 2.51 \text{ for Na} \\ &= 1.76 \text{ for Rb} \\ &= 3.67 \text{ for Cs} \\ &= 5.33 \text{ for Tl.} \end{aligned}$$

Fermi, using his statistical atom model and numerical calculation of the eigenfunctions, has given the values 4.11 and 1.89 for Na and Cs respectively.

[9] The analogous approximation can always be made when k is negative. For example, we obtain

$$\overline{r^{-4}} = \frac{\frac{3}{2}Z^2 - \frac{1}{2}(l+\frac{1}{2})^2E}{(l+\frac{1}{2})^5z\epsilon^{-3/2}} \cong \frac{3}{2}\,\frac{Z^2z^2}{n^3_{eff}(l+\frac{1}{2})^5}$$

Of course this like (10) will hold only for $l>0$ since for $l=0$, $\overline{r^{-3}}$ and $\overline{r^{-4}}$ diverge. For positive values of k similar approximations may be made giving, for example

$$\overline{r^2} = \frac{1}{2}\,\frac{n^2_{eff}}{z^2}\left\{5n^2_{eff} - 3\left(l + \frac{1}{2}\right)^2\right\}.$$

[10] E. Schrodinger, Zeits. f. Physik **4**, 347 (1921).

[11] S. Goudsmit and L. A. Young, Nature **125**. 461 (1930).

It is unfortunate that it is difficult to give a criterion for the accuracy of these results. One can hardly expect them to be correct except as to order of magnitude.

Appendix

Mean Value Integrals

In calculating the mean values $\overline{r^k}$ for hydrogenic atoms the following integrals were encountered:

$$I_k^{(1)} = \tfrac{1}{4}\int_0^{r_1} r^k Q^{-1} \exp\left\{-2\int_r^{r_1} Q dr\right\} dr$$

$$C_k = \tfrac{1}{2}\int_{r_1}^{r_2} r^k Q^{-1} dr$$

$$U_k = \tfrac{1}{2}\int_r^{r_2} r^k Q^{-1} \sin\left\{2\int_{r_1}^{r} Q dr\right\} dr$$

$$I_k^{(3)} = \tfrac{1}{4}\int_{r_2}^{\infty} r^k Q^{-1} \exp\left\{-2\int_{r_2}^{r} Q dr\right\} dr.$$

The integral C_k was evaluated exactly, but it was necessary to introduce simplifying assumptions in evaluating the correction integrals $I_k^{(1)}$, U_k, and $I_k^{(3)}$. It was found that

$$C_k = \tfrac{1}{2}\pi\epsilon^{-1/2} P^{(k+1)/2} P_{-(k+2)}(\tfrac{1}{2}SP^{-1/2}) \qquad k < -1$$
$$= \tfrac{1}{2}\pi\epsilon^{-1/2} P^{(k+1)/2} P_{k+1}(\tfrac{1}{2}SP^{-1/2}) \qquad k \geqq -1$$

where P is the product of the roots of Q^2, r_1 and r_2, and $P_m(x)$ is the m-th Legendre polynomial of the argument x. No general formula was found for the integral $U_{(k)}$ but after simplifying approximations had been made it was possible to evaluate the integrals for special values of k with the results:

$$U_2 \cong \frac{(-1)^{n_r}}{16}\pi\epsilon^{-1/2}\left[3D^2S\frac{J_1(\xi)}{\xi} + S^3 J_0(\xi)\right]$$

$$U_{-2} \cong (-1)^{n_r}\pi\epsilon^{-1/2}S^{-1}\left[J_0(\xi) + \left(\frac{D}{S}\right)^2\frac{J_1(\xi)}{\xi} + 3\left(\frac{D}{S}\right)^4\frac{J_2(\xi)}{\xi^2} + \cdots\right].$$

The values of U_k for other values of k were obtained by use of the recurrence relation

$$U_{k-1} = \frac{2}{k+1}\frac{dU_k}{dS}.$$

In these relations $S = r_1 + r_2$, $D = r_2 - r_1$, $\xi = (n_r + \tfrac{1}{2})\pi$, and $J_n(\xi)$ is the n-th Bessel function of the argument ξ. The integrals $I_k^{(1)}$ were found to be given approximately by:

$$I_k^{(1)} \cong \frac{(l+\tfrac{1}{2})^{2k+2}\exp\{l+\tfrac{1}{2}\}}{2^{k+1}(2l+k+5)}.$$

Finally, for the integrals $I_k^{(3)}$ we find

$$I_2^{(3)} = \frac{1}{8} r_2^3 \epsilon^{-1/2} \exp\left(\frac{\beta}{2}\right)\left[\left(1+\frac{1}{2\beta}\right)K_0\left(\frac{\beta}{2}\right)+\left(1+\frac{3}{2\beta}+\frac{2}{\beta^2}\right)K_1\left(\frac{\beta}{2}\right)\right]$$

$$I_1^{(3)} = \frac{1}{8} r_2^2 \epsilon^{-1/2} \exp\left(\frac{\beta}{2}\right)\left[K_0\left(\frac{\beta}{2}\right)+\frac{\beta+1}{\beta}K_1\left(\frac{\beta}{2}\right)\right]$$

$$I_0^{(3)} = \frac{1}{8} r_2 \epsilon^{-1/2} \exp\left(\frac{\beta}{2}\right)\left[K_0\left(\frac{\beta}{2}\right)+K_1\left(\frac{\beta}{2}\right)\right]$$

$$I_{-1}^{(3)} = \frac{1}{4} \epsilon^{-1/2} \exp\left(\frac{\beta}{2}\right)K_0\left(\frac{\beta}{2}\right)$$

$$I_{-2}^{(3)} = \frac{1}{4} \beta r_2^{-1} \epsilon^{-1/2} \exp\left(\frac{\beta}{2}\right)\left[K_1\left(\frac{\beta}{2}\right)-K_0\left(\frac{\beta}{2}\right)\right]$$

$$I_{-3}^{(3)} = \frac{1}{6} \beta r_2^{-2} \epsilon^{-1/2} \exp\left(\frac{\beta}{2}\right)\left[\beta K_0\left(\frac{\beta}{2}\right)-(\beta-1)K_1\left(\frac{\beta}{2}\right)\right]$$

where $\beta = 2\epsilon^{-\frac{1}{2}} r_2$ and $K_\nu(x)$ is the Bessel function of the third kind of order ν.

ON THE ORIGIN OF THE BANDS IN THE SPECTRUM OF MERCURY VAPOR

(Remarks on a Paper by R. Rollefson)

By S. Mrozowski

Physical Laboratory, Society of Sciences and Letters, Warsaw

(Received July 28, 1930)

Abstract

Arguments of Rollefson, admitting the existence of molecules of small energy of binding in mercury vapor, are criticized. It is shown that the experimental facts lead to the following conclusions: (1) That absorption is due to the molecules which exist before irradiation of the vapor; (2) probably all absorption bands have the same origin; (3) the bands are due to molecules with great energy of binding (17 kcal/mol); and (4) the emission bands lying in the region of the absorption bands are emitted by the molecule in returning to its normal electronic level.

A new observation is made of shaded absorption bands at 2267 and 2247A.

IN A very interesting paper[1] Rollefson tries to reconcile the contradictory results given by Franck and Grotrian[2] and Koernicke[3] and those which the author has obtained.[4] Rollefson concludes that different bands in mercury vapor spectrum have different origin, particularly the band 2540A may be accounted for the presence of Hg_2 molecules having small energy of binding (1.4 kcal/mol). The existence of such molecules seems very probable from the theoretical standpoint. Leaving aside the theoretical discussions the author will try in this paper to consider critically the arguments of Rollefson only from the point of view of experimental facts. In this connection the author would like to point out a series of very simple experimental facts, which were generally not taken into account in building the hypotheses relating to the origin of the band spectrum of mercury vapor.

We have to remark that the contradiction pointed out by Rollefson does not exist, for the author has shown in a former paper,[3] that Koernicke, (and probably also Franck and Grotrian, whose preliminary measurements were improved by Koernicke) applied an erroneous formula and therefore the results cannot be convincing. Up to the time of a precise repetition of the measurements of Koernicke, there is no need to search for new points of view which would reconcile the results of Koernicke and those of the author.

To recall the complicated structure of the absorption bands of mercury vapor the graph of this band spectrum is given in Fig. 1.

The band series a[5] and e were found by Rayleigh, b and c is the fine struc-

[1] Rollefson, Phys. Rev. **35,** 1177 (1930).

[2] Franck and Grotrian, Zeits. f. techn. Physik **3,** 194 (1922).

[3] Koernicke, Zeits. f. Physik **33,** 219 (1925).

[4] Mrozowski, Zeits. f. Physik **55,** 338 (1929).

[5] The fact that these bands can be due to Hg_2O molecules (Walter and Barrat, Proc. Roy. Soc. **A122,** 201, 1929) does not influence the consideration in this paper.

ture of the 2540A band discovered by Wood and Voss,[6] *d* the bands observed long ago by Wood, finally the two shaded bands *f* and *g* were found by the author on the photographs of absorption bands. The absorption tube in these experiments had a length of 40 cm. A hydrogen quartz lamp wholly immersed in water, with electrodes also cooled by running water, was used

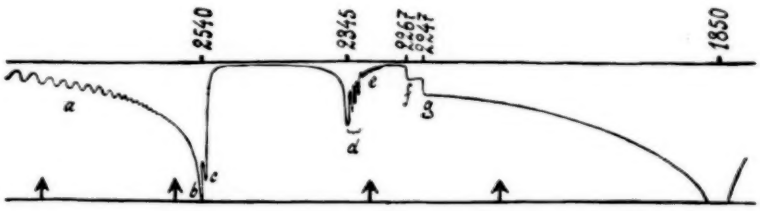


Fig. 1.

as the source of the continuous spectrum. (At 3000 volts a current of 2 amp. could be maintained, but in the experiments 0.5 amp. was mainly used.) The shaded bands *f* and *g* are represented in Fig. 2 where we have

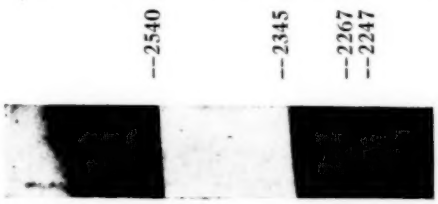


Fig. 2.

a photograph taken by a small Liess spectrograph type A.

Different hypotheses were made about the origin of the mercury absorption bands. One of these was, that the molecules which are responsible for the absorption are formed from excited and unexcited atoms during the irradiation. But the fact that the spectral lines of other elements separated by the monochromatic illuminator are absorbed (and excite the band emission) is in contradiction with the former hypothesis. In general the experiments show conclusively, that: (1.) The absorption is due to the molecules which exist before the irradiaton of the vapor, because the form of the absorption bands does not depend in wide limits upon the intensity of the light transmitted through the vapor (variations of time of exposure from $\frac{1}{2}$ min. to 24 hours were used by the author). Nor do these bands depend upon the distillation of vapor (Niewodniczański).[7]

The theory of quasimolecules of Born and Franck and the theory of Oldenberg (interaction of radiation and impact) which is identical with it can only explain the widening of the atomic absorption lines which appears in dense vapor, but they cannot explain the complicated structure of the band

[6] Wood and Voss, Proc. Roy. Soc. **A119,** 698 (1928).

[7] Niewodniczański, Compt. Rend. Soc. Pol. de Physique **3,** 31 (1927).

absorption spectrum of mercury vapor. There remain two possibilities: molecules with small or great energy of binding. In the first place it is to be proved, that: (2.) Probably all the absorption bands of mercury vapor have the same origin. This can be seen on the basis of the former measurements of dependence of the band absorption coefficient upon the temperature of saturated mercury vapor, made for wave-lengths 2749, 2573 and 2144A[4] and also recent measurements made with a quite different method for wave-lengths 2749, 2573, and 2313A[8] by the author. All these experiments show a good proportionality between the dependences of the absorption coefficients for four different wave-lengths lying in different absorption regions (corresponding spots are marked by arrows in Fig. 1). These measurements show further that: (3.) The bands are due to atoms or molecules with great energy of binding (17 kcal/mol). The heat of evaporation found from the change of the absorption coefficients is, on the average, 12 kcal/mol; we can say that the absorption is due to atoms (the heat of evaporation $\lambda_1 = 14.5$ kcal/mol; the difference could be explained by experimental errors) or molecules with energy of binding approximately of 17kcal/mol($= 2\lambda_1 - 12$kcal/mol) The suggestion of Rollefson, that in the case of absorption of the line 2573A two different phenomena overlap: the absorption of molecules of small energy of binding and of quasimolecules (relatively the interaction of radiation and impact) does not evidently suffice to explain the experimentally shown proportionality of absorption coefficient for the wave-length 2573A to the concentration of atoms (approximately), for the phenomena mentioned are proportional to the square of density of vapor (heat of evaporation of $2\lambda_1$ also), which is too far outside the limits of the experimental errors of the methods applied.

Considering the emission bands of mercury vapor, we can refer to the graph shown in a recent paper of the author[9] (Fig. 4 of the paper; the only uncertainty relating to the emission in the 2000–1850A interval in fluorescence was explained recently by Eliashewich and Teremin[10] in agreement with the predictions of the author). Contrary to the suggestion of Rollefson, we have to assume, that (4.) The emission bands lying in the regions of the absorption bands (therefore excluding only the continuous band 4850A) are emitted by the molecules returning to their normal electronic level. That is shown by the proportionality of the intensity of fluorescence to the intensity of the exciting light (Wood and Voss[6]) and the fact, that for monochromatic excitation even of very weak intensity (time of exposure on a small spectrograph 100 hours) there appears the anti-Stokes part of the fluorescence band. For example, the band of Steubing (2350-2000A, the author has previously shown that they belong to the pure vapor of mercury[4]) is emitted by the transition to the normal electronic level and shows a series of separated oscillation bands. Therefore we can affirm that the molecules in the normal and in the excited states alike ought to have a relatively great energy of

[8] Mrozowski, Wszechswiat **3**, 101 (1930).
[9] Mrozowski, Zeits. f. Physik **62**, 314 (1930).
[10] Eliashewich and Terenin, Nature **125**, 856 (1930).

dissociation. That is in agreement with conclusion 3, for from the modern point of view about the structure of the atomic spectra we have to omit the first case of point 3. Moreover we can remark, that by assuming with Rollefson that the band 2540A is due to molecules of small energy of binding, it should be difficult to explain the structure of that band observed in emission by Takeo Hori.[11]

Although the correlation of molecular and atomic electron levels given previously by the author[4] may be considered only as provisional, it seems to us however, that the result given by the experiments, namely that there are molecules of mercury vapor with great energy of dissociation, is a very important one. We have to bear in mind that in the first place there is a contradiction with the results of measurements of the mercury vapor density on the one hand and with the theoretical view on the possibility of formation of such molecules from atoms with closed electronic shells on the other hand. The explanation of these contradictions can be expected from new precise measurements.

[11] Takeo Hori, Zeits. f. Physik **61**, 481 (1930).

ABNORMAL SHOT EFFECT OF IONS OF TUNGSTOUS AND TUNGSTIC OXIDE

By John S. Donal Jr.
University of Michigan

(Received August 18, 1930)

Abstract

Part I. The formation of positive ions of tungstous and tungstic oxide. It was found that when oxygen attacks a hot tungsten filament, positive ions are formed, and a mass spectrograph test has shown these ions to be a mixture of tungstous and tungstic oxide. The positive ion current is proportional to the pressure of the gas, as is the rate of formation of the oxide. The evaporation of an oxide layer following the pumping away of the oxygen, yields an ion current which falls away exponentially with the time. The ratio of the number of charged particles of the oxide, to the number of uncharged particles, is approximately one in twenty thousand, and this ratio does not vary greatly with the temperature of the tungsten filament.

Abnormal shot effect of the currents of positive ions. When the positive ions are drawn to a collecting electrode, in series with which there is a tuned shot circuit, an abnormal shot voltage is developed. The hypothesis is advanced that this shot voltage arises from fluctuations in the inter-electrode capacitance of the tube, caused by variations in the thickness of the electron sheath surrounding the anode as the sheath is penetrated by the positive ions. On this basis, an expression for the resulting mean square shot voltage is presented. This expression was found to be correct as regards the variation of the shot voltage with the ion current, and it has yielded reasonable values for the change in the inter-electrode capacitance.

Part II. Abnormal shot effect due to the trapping of the positive ions. When positive ions of tungstous and tungstic oxide were trapped in the minimum of potential surrounding a hot cathode, an abnormal shot voltage resulted. This abnormality was found to vary inversely with the square of the resonance frequency of the tuned shot circuit, analogously with the "flicker-effect." An expression for the mean square shot voltage has been developed, which permits the calculation of the average plate current increase due to the trapping of a single ion. The simultaneous measurement of the shot voltage and the total plate current increase has permitted the determination of the number of electrons released by each positive ion, and also the average time of trapping of the ions. From a knowledge of the trapping time, values of the electron space charge density in the minimum of potential were computed. These values were in agreement with those computed from a knowledge of the filament temperature, and the space current, and the two methods of arriving at the space charge density yielded magnitudes which varied in the same manner with the electron current through the tube.

Introduction

IT IS well known, particularly from the work of Kingdon and Charlton,[1] that positive ions produced at the surface of a hot filament may become trapped in the minimum of potential adjacent to the cathode surface, and will in that case cause a partial break-up of the space charge and a release of electrons to the anode. It was felt that the fluctuations of the plate current

[1] K. H. Kingdon and E. E. Charlton, Phys. Rev. **33**, 1011 (1929).

due to this trapping process would develop an abnormal mean square shot voltage across a tuned shot circuit, and that from measurements of this voltage much information concerning the process could be obtained.

It was proposed to make use of the fact that caesium vapor becomes ionized upon contact with a hot tungsten filament, but in the course of some other work it was found that positive ions are produced when oxygen attacks a hot tungsten surface, and this source of ions was used in the experiments. Part I will be devoted to a brief discussion of this phenomenon, and of the experimental conditions governing the formation of the ions.

Before undertaking the work to be described, measurements were made of the charge carried by electrons in temperature limited currents from the filaments of several commercial vacuum tubes. These determinations were carried out with the idea of placing the apparatus in correct operating condition, use being made of the known charge carried by the electron.

As a preliminary piece of work, a determination was made of the charge of electrons produced by the ionization of argon gas. In the course of this work, large abnormalities were found to be present, as long as the electron currents were space charge limited. A detailed study of these effects was made, but space will not permit further discussion here. It will suffice to say that, when the abnormalities were eliminated, the electrons produced by the gas ionization were found to carry the accepted value of electronic charge, within the probable error of the experimental work.

The shot effect apparatus used in these determinations was of essentially the same design as that employed by Williams and Vincent,[2] in previous researches on the measurement of the charge of the electron. The experimental tubes were assembled by the writer, and varied considerably in design. The vacuum system employed was capable of maintaining a pressure of 10^{-7} mm of mercury, and was equipped with the means of introducing and mixing various gases.

Part I. The Formation of Positive Ions When a Hot Tungsten Filament is Attacked by Oxygen Gas

It is well known, from the work of Langmuir and others, that hot tungsten is attacked by oxygen gas, with the formation of tungstic oxide. The rate of attack is proportional to the oxygen pressure, if the filament is maintained at a constant temperature. When the metallic surface is above about 1250°K, the oxide is thrown off, but when the surface temperature is below this value, the oxide forms a thick layer and the filament is very perceptibly darkened. If the coating is permitted to form on the metal, the rate of attack by the gas is decreased.

In the course of some experiments on the abnormal shot effect of currents of positive ions from pure tungsten filaments, the ions were accelerated through gases. It was found that when oxygen was admitted to the system, the positive ion current was greatly increased, and that the increase was directly due to the oxygen itself. It was immediately surmised that the ions

[2] N. H. Williams and H. B. Vincent, Phys. Rev. **28,** 1253 (1926).

were those of tungstic oxide, with the possibility of tungstous oxide being present also. Langmuir,[3] in one of his earlier papers, has shown that the attack of a tungsten surface by oxygen must take place in two steps, and he has suggested that the first may well be the formation of tungstous oxide on the metallic surface, oxygen further uniting with this compound to form the common trioxide.

As would be expected, the number of positive ions formed is very much a function of the gas pressure, the filament temperature, and the area of the tungsten surface. An idea of the magnitude of the positive ion current may be had from the fact that with a tungsten surface 1 cm² in area, held at a temperature of 2300°K, one may expect the saturation value of the positive

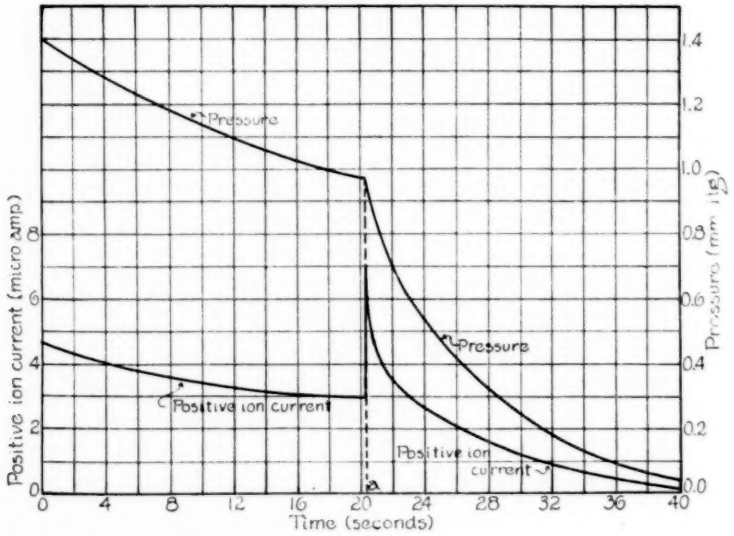


Fig. 1. Behavior of positive ion current and gas pressure, before and after evaporation of oxide layer.

ion current to be of the order of 5×10^{-5} amperes, with an oxygen pressure of 1 mm of mercury. This current is quite naturally difficult to maintain, since the gas is removed exceedingly rapidly by the filament, which is in turn eaten away. These facts have made the determination of the exact experimental relationships difficult and in some cases impossible.

As the result of experiments made with over one hundred tubes, it has been found that the rate of formation of the positive particles is directly proportional to the amount of oxygen entering into chemical combination with the tungsten to form the oxide, and it was upon this fact that the assumption as to the nature of the ions was based.

The behavior of the positive ion current and the gas pressure, with the time, is illustrated in Fig. 1. To the left of the point (*a*) on the time axis the filament temperature was so low that the oxide remained on the metallic surface. The positive ion current and the rate of decrease of gas pressure

[3] I. Langmuir, Proc. Am. Chem. Soc. **37**, 1139 (1915).

were therefore low. At the arbitrary point (*a*) in time the filament temperature was increased sufficiently to break away the oxide layer and to prevent its re-formation. It will be noted that the positive ion current increased sharply and then fell away rapidly, as did the gas pressure, since the rate of clean-up by the fresh metallic surface was very much greater and the temperature had been increased. It will be seen from the figure that the ion current was proportional to the gas pressure, at the points beyond (*a*) in time. The rate of formation of tungstic oxide is already known to be proportional to the gas pressure, and we thus see the direct relationship between the rates of formation of the ions and of the oxide. Subsequent determinations of the variation of the ion current with the oxygen pressure have shown the direct

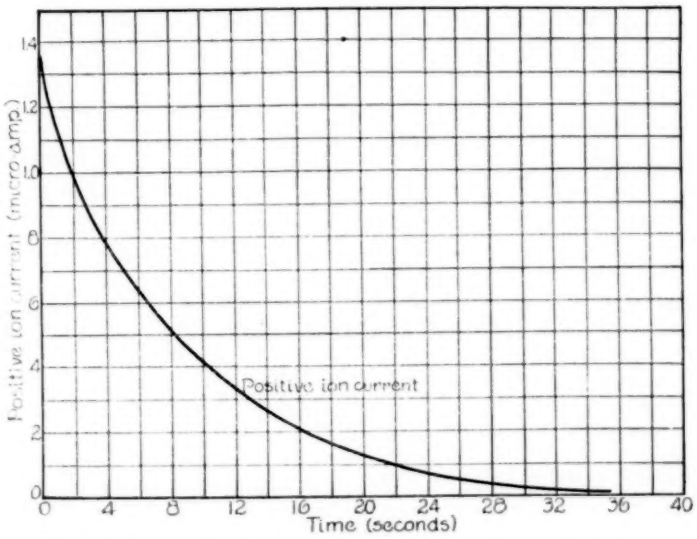


Fig. 2. Positive ion current obtained during the evaporation of an oxide coating from a tungsten filament.

proportionality to hold rigidly when the other experimental conditions are maintained constant.

It was naturally supposed that the ionization process takes place either at the moment of formation of the oxide or at the instant when the particle leaves the metallic surface, and that once the oxide is formed the oxygen gas is no longer essential. In order to test this assumption, an oxide layer was formed on a relatively cool filament, the remaining gas then pumped away and the tube walls well out-gassed. Following this treatment the filament temperature was increased sharply to a value slightly above that which would just permit the adherence of the coating. The variation of the ion current with time was then determined and the result is illustrated in Fig. 2.

From the shape of the curve of Fig. 2, it is apparent that the number of charged particules leaving the filament was proportional to the amount of oxygen resident there. No appreciable gas was given off during this breakdown process.

At the writer's request, L. P. Smith and L. Barnes, of Cornell University, lately have been so kind as to make a mass spectrograph determination of the nature of the positive ions under discussion. Their results are illustrated in Fig. 3.

It will be noted from the figure that both the singly charged dioxide and trioxide ions were present, in the approximate ratio of one to three, respectively. The presence of the dioxide ions is particularly interesting, in the light of Langmuir's conclusions as to the nature of the two-step process involved in the oxide formation.

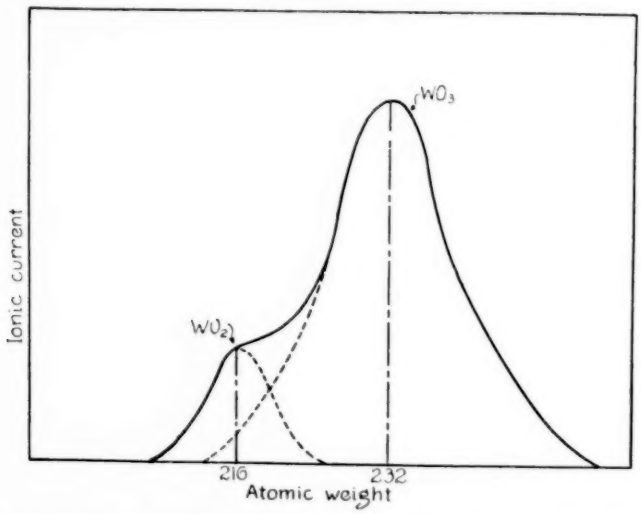


Fig. 3. Mass-spectrograph determination of the nature of the ions formed when oxygen attacks a hot tungsten filament.

The Relative Number of Oxide Particles Which Leave the Surface Carrying A Charge

Experiments were undertaken to determine what proportion of the particles of the oxide actually becomes charged, and how this proportion varies with the temperature of the metallic filament. In order to accomplish this, a tungsten filament was heated in an atmosphere of oxygen of a known initial pressure and volume. The total number of removed molecules of oxygen was computed from the change in pressure during the period of operation of the filament, and from this was determined the number of molecules of either WO_2 or WO_3 formed. During the heating of the filament the saturation value of the positive ion current to the cathode was measured, and the graphical integration of this curve yielded the number of ionized molecules of oxide.

The results of several representative determinations are presented in Table I. Since WO_2 is unstable chemically, the values have been computed on the basis of WO_3 alone. It is of course possible that all of the WO_2 leaving the filament was charged, and in fact it may have left the filament only by virtue of its charge.

The runs of Table I were made at different filament temperatures. The results, as regards the variation of the numbers of charged particles with the temperature, are particularly interesting. One would perhaps expect the oxide molecules to become charged more readily at the higher temperatures. This is obviously not the case, the trend being decidedly in the opposite direction toward the end of the table. It must be pointed out, however, that the smaller relative numbers of charged particles at the higher temperatures may be spurious. With the exceedingly hot filaments used in the latter runs of the table, the electron emission was high, from actual measurements, and there was a dense electron space charge surrounding the filament, which may well have brought about the removal of many of the oxide ions by recombination.

TABLE I. *Values of the ratio of the number of uncharged particles of WO_3, to the number of charged particles.*

Run	Approximate filament temperature	M/M_+
4,5	Below oxide evaporation point	14840
6	" " " "	15580
2	Just above oxide evaporation point	14940
13	" " " " "	14920
14	" " " " "	13800
10	Bright yellow heat	17560
15	" " "	17220
16	" " "	16080
17	" " "	15960
19	White heat	20600
20	" "	19940
22	" "	20560
23	" "	22600
1	Just below burn-out temperature	40400
8	" " " "	38560
9	" " " "	69200

We may conclude that, in spite of the great variation in gas pressure, filament surface condition and surface temperature from run to run, the ratio of the number of charged particles to the number of uncharged particles was of the order of one to twenty thousand and that, as regards temperature in particular, no marked variation was found.

It may be of interest to know that in experiments in which the temperature was increased uniformly until the filament burned out, the positive ion current almost invariable decreased slightly at the temperatures just below the fusion point of the tungsten. From the other results, this indicates a decrease in the rate of formation of the oxide at the highest temperatures, a result which has been predicted by Langmuir.[4]

ABNORMAL SHOT EFFECT CAUSED BY THE POSITIVE IONS OF WO_2 AND WO_3

A brief discussion of a series of determinations of the shot voltage developed by the currents of positive ions will now be presented. A tuned shot circuit was placed in the lead to the collecting electrode, and it was found that when the positive ion current traversed this circuit the resulting mean square

[4] I. Langmuir, Proc. Am. Chem. Soc. **37**, 1139 (1915).

shot voltage was abnormal by a factor as great as one thousand, depending upon the experimental conditions.

It was apparent that the magnitude of the shot voltage was very greatly dependent upon the emission of electrons from the hot anode, entirely aside from the rate of attack by oxygen. In short, not only the positive ions were necessary to develop the abnormal shot voltage, but an electron space charge was necessary as well. This fact is clearly illustrated in the plots of Fig. 4, which are representative of many such runs.

Fig. 4 illustrates the variation of the positive ion current, the abnormal mean square shot voltage and the electron emission (determined subsequently with duplicated expermental conditions as to pressure and tempera-

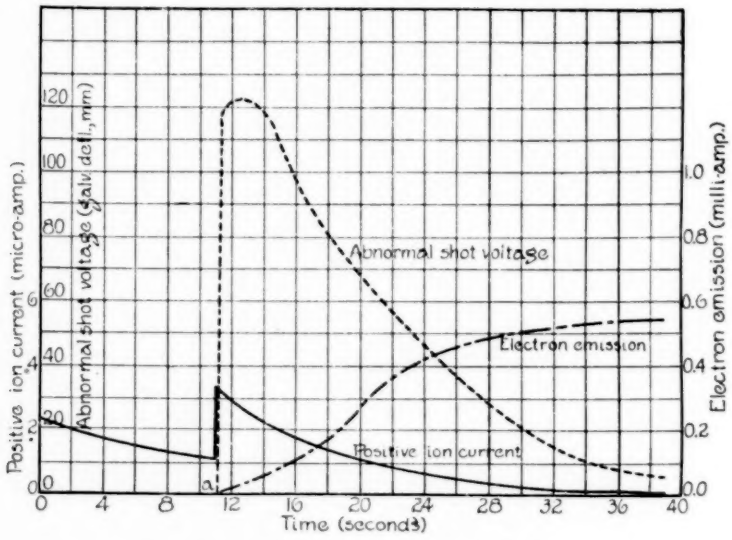


Fig. 4. Behavior of abnormal shot voltage and electron emission, following evaporation of oxide layer.

ture) as functions of the time. Again at the arbitrary time (*a*) the filament was heated sufficiently to break away the oxide layer. It will be noted that before the oxide was evaporated there was a current of positive ions, as in Fig. 1, but there was no electron emission nor was there any abnormal shot voltage. With the ion currents used, the normal shot voltage was so small as to be unmeasurable. As soon as the oxide layer was removed, however, a shot voltage, abnormal by a factor of approximately 1000, appeared, whereas the positive ion current increased by only a factor of two. It is to be particularly noted that the electron emission began to build up as soon as the layer was evaporated, increasing rapidly as the oxygen was removed from the tube. A large number of determinations were made in order to check this point, and in all cases it was found that the positive ion current alone was a necessary but not sufficient condition for the abnormal shot voltage fluctuations, the electron space charge being essential also.

In order to explain the observed mean square shot voltage, several hypotheses were considered. It will be remembered that the field used in these experiments was a retarding one for electrons, which therefore could only leave the hot anode by virtue of their emission velocities. This resulted in an equilibrium being established between the electrons leaving the filament and those returning, the only electrons present in the tube being those in the dense sheath close to the anode surface. The only direct current traversing the shot circuit was the positive ion current, and this had been found insufficient to yield the abnormality without the presence of the electron space charge.

Of the hypotheses advanced to explain the abnormality, all but one were found to be untenable, due to the peculiar experimental conditions. The theory adopted was that the abnormality arose from fluctuations in the inter-electrode capacitance of the tube itself, caused by variations in the size of the electron space charge sheath about the hot anode as the sheath was penetrated by the massive positive particles of oxide.

On the basis of the above assumption, the writer has developed an expression for the mean square shot voltage to be found across a tuned shot circuit, in terms of the number of positive particles formed per second and the average change in the inter-electrode capacitance due to the passage of a single ion. For this expression we have:

$$\overline{E}^2 = \frac{2NK^2V^2\overline{\Delta C}^2L^2A_0^2A}{R^2C^2}. \tag{1}$$

Here, N is the number of positive ions formed per second, K a constant depending upon the time constant of the circuit, V the negative plate potential in volts, $\overline{\Delta C}$ the average change in the inter-electrode capacitance due to the passage of a single ion, L, R, C are the constants of the tuned shot circuit, A_0 the voltage amplification of the amplifier at resonance frequency and A is the area under the amplifier resonance curve, as plotted against the frequency.

By comparing the $\overline{E}^2$ of Eq. (1) with the normal shot effect $\overline{E}^2$, as developed in a tube known to yield the accepted value of the electronic charge it was found possible to measure $\overline{\Delta C}$ without a knowledge of the circuit constants or the characteristics of the amplifier.

Under the experimental conditions, Eq. (1) could only be subjected to direct test as regards the variation of $\overline{E}^2$ with N. The form of this variation was determined, and the result was entirely in accordance with the equation. Also, from measurements of $\overline{E}^2$, values of $\overline{\Delta C}$ were computed, and these were found to be of the order of 10^{-4} times the actual electrode capacitance of the tube. Since, if Eq. (1) were based upon wrong assumptions, we might just as readily obtain values for $\overline{\Delta C}$ greater than the electrode capacitance itself, the computed results are thought to lend considerable support to the assumption as to the nature of the abnormality.

Part II. Abnormal Shot Effect Due to the Trapping of Positive Ions in the Minimum of Potential

Introductory. This portion of the investigation was undertaken in order to test the writer's prediction that if positive ions were to become trapped in the minimum of potential surrounding a hot cathode, an abnormal shot effect would result. One would expect the ion, when trapped, to alter the potential distribution in the space charge region, and to cause an increase in the current to the plate, which increase should persist for the duration of the stay of the ion in the region surrounding the cathode. This change in the current would not be constant, of course, but would itself fluctuate as the ion oscillated about its mean position. The phenomenon would be fundamentally different in its effect upon the shot circuit, however, from the simple passage of an ion across the tube. In the latter case the duration of the event would certainly be less than the period of the tuned shot circuit used in these experiments (10^{-5} seconds). If the ion were to become trapped, however, and if neither the loss of ions by escape from the ends of the electrodes nor the loss by recombination with electrons were excessive, the duration of the event might be expected to be greater than the period of the tuned shot circuit. As has been pointed out by W. Schottky,[5] the theoretical attack for these two cases is fundamentally different.

It was further hoped that the method of attack evolved by Schottky could be made applicable to the abnormal shot effect resulting from the trapping of the ions, and that, on this basis, a simple experimental procedure could be developed for determining the average effect of the positive ions under different conditions.

Finally, the writer wished to find a means of checking both the applicability of the Schottky theory and the quantitative experimental results to be obtained.

Preliminary experiments. In the course of the work with argon, mentioned in the Introduction, it had been found that the positive ions formed by the ionization of the gas yielded a very abnormal mean square shot voltage. This was due to the release of groups of electrons from the space charge region, as a result of the low mobility of the ions in comparison with the electron mobility. This phenomenon did not lend itself to quantitative measurements, however, since the ions were dependent upon the field for their formation, the ionization causing great instability in the tube plate characteristic and erratic readings as a consequence. Furthermore, there was no dependable method of measuring the actual number of ions formed per second.

When use was made of the ions formed by contact with oxygen with the hot tungsten filament, however, the phenomenon was fully under control. The number of ions formed was not a function of the field, but, the filament temperature being constant, was governed entirely by the oxygen pressure. This yielded separate control of the number of ions and of the space charge

[5] W. Schottky, Phys. Rev. **28**, 74 (1926).

density, and the number of ions formed per second could be measured at once by simply reversing the potential of the collecting electrode.

With the use of oxygen, and when the electron currents were space charge limited, it was found that the mean square shot voltage was abnormal by a factor of from 100 to 1000, depending upon the density of the space charge. The abnormality disappeared as soon as the electron currents were temperature limited. It remained to be determined whether or not the large shot voltage was due to the trapping of the ions, giving a result of the same nature as the "flicker effect."

Schottky[6] has shown that, if the abnormal shot effect is caused by a primary event whose duration is greater than the period of the tuned shot circuit used, we would expect the shot voltage to vary inversely as the square of the resonance frequency of the tuned circuit. The particular case for which Schottky's theory was developed was the shot effect found experimentally by J. B. Johnson,[7] and explained on the basis of fluctuations in emission caused by the temporary residence of foreign molecules on the emitting surface. This phenomenon followed the inverse square variation with the frequency of the tuned shot circuit.

Determinations were made of the trend with frequency of the abnormality in argon, and of the abnormality resulting from the positive ions of tungstous and tungstic oxide. In the former case, no definite frequency variation was found in the region from 100 to 250 kilocycles. In the latter case, however, the shot voltage was found to exhibit, within a few percent, the inverse square frequency variation. This at once classified the abnormality in oxygen as a phenomenon of the nature of the "flicker effect," with a primary event lasting longer than the period of the tuned shot circuit, and indicated that the cause of the abnormality was undoubtedly the trapping of the positive ions in the region of the minimum of potential, rather than the simple passage of an ion through the space charge region.

Theory. If we let γ be the average plate current increase due to the trapping of a single positive ion, we find that it is possible to obtain an expression for γ in terms of quantities easily measured experimentally. It may be pointed out that for the moment we are not interested in the factors limiting the duration of this average current increase, but only in the fact that, as was determined by the experimental frequency trend, the duration of the increase is substantially greater than the period of the tuned shot circuit.

The shot circuit consists of a tuned impedance in the plate lead of a two electrode tube and, in addition to the steady average value of the plate current, we have flowing through it a fluctuation current (due to electrons released from the space charge region) which may be written in terms of its Fourier components as follows:

$$j = \sum_{k=1}^{\infty} A_k \cos \omega_k t + \sum_{k=1}^{\infty} B_k \sin \omega_k t = \sum_{k=1}^{\infty} C_k \sin (\omega_k t + \Phi_k) \qquad (2)$$

[6] W. Schottky, Phys. Rev. **28,** 74 (1926).

[7] J. B. Johnson, Phys. Rev. **26,** (July, 1925).

where A_k, B_k, and C_k are partial current amplitudes and ω_k is $2\pi x$ the frequency of the kth component.

Of course:

$$\left.\begin{aligned} \overline{A_k^2} + \overline{B_k^2} &= C_k^2 \\ \text{and:} \qquad \overline{A_k^2} &= \overline{B_k^2} \end{aligned}\right\} \tag{3}$$

Schottky[8] has evaluated these coefficients in his treatment of the "flicker effect," with the only difference that the fluctuations in the electron current originated at the cathode surface, rather than at the potential minimum, as in our case. Schottky obtained the result:

$$\overline{A_k^2} = \overline{B_k^2} = \frac{2}{T} \cdot \frac{\overline{\beta^2}\gamma^2\alpha}{\alpha^2 + \omega_k^2} = \frac{2}{T} \cdot \frac{n\gamma^2}{\alpha^2 + \omega_k^2} \tag{4}$$

(since $\overline{\beta^2} = N_0$ from probability considerations, and $N_0\alpha = n$.)

In this expression T is a time of observation, α the reciprocal of the time of trapping of an ion, N_0 the average number of ions in the space charge region at any instant, β the momentary deviation from this average, n the number of ions formed per second and ω_k is $2\pi x$ the frequency of the kth component of the fluctuation current.

We therefore have:

$$\overline{C_k^2} = \frac{4}{T} \cdot \frac{n\gamma^2}{\alpha^2 + \omega_k^2} \cdot \tag{5}$$

If a fluctuation current j is passed through a tuned shot circuit, and the resulting voltage fluctuations across the terminals of the tuned shot circuit are amplified, Williams and Vincent[9] have shown that the mean square shot voltage is given by:

$$\overline{E^2} = \sum_k \overline{E_k^2} = \frac{L^2A_0^2}{2R^2C^2} \sum_{k=1}^{\infty} C_k^2\Phi(x) \tag{6}$$

where A_0 is the voltage amplification of the amplifier at resonance frequency and $\Phi(x)$ is the curve obtained by plotting the relative power output at different frequencies against $x = \omega_k/\omega_0$.

If $dx = 2\pi/\omega_0 T$ be substituted in (6), together with C_k^2 from Eq. (5), we obtain:

$$\overline{E^2} = \frac{1}{\pi} \cdot \frac{L^2A_0^2}{R^2C^2\omega_0} \int_0^{\infty} \frac{n\gamma^2}{x^2 + \alpha^2/\omega_0^2} \Phi(x)dx. \tag{7}$$

The integrand of Eq. (7) may be considerably simplified if $\alpha << \omega_0$. In the experimental case, α will be found to be of the order of 10^4 and ω_0 was approximately 10^6. Therefore α^2/ω_0^2 was approximately 10^{-4} and could be neglected in comparison with x^2, which was very nearly unity.

Eq. (7) then becomes:

[8] References 5 and 6.

[9] N. H. Williams and H. B. Vincent, Phys. Rev. **28**, 1253 (1926).

$$\overline{E^2} = \frac{n\gamma^2 L^2 A_0^2}{R^2 C^2 \omega_0} \int_0^\infty \frac{\Phi(x)}{x^2} dx. \tag{8}$$

A further simplification may also be carried out. If we write $1/z$ for x in the integrand of (8), we obtain:

$$\overline{E^2} = \frac{n\gamma^2 L^2 A_0^2}{R^2 C^2 \omega_0} \int_0^\infty \Phi\left(\frac{1}{z}\right) dz. \tag{9}$$

But it may be shown that if the amplifier resonance curve is sharp and symmetrical, as is true in the experimental case, we may write:

$$z = 1/z$$

with an error of only about 0.1%.

Thus:

$$\Phi\left(\frac{1}{z}\right) \leqq \Phi(z)$$

and (9) becomes:

$$\overline{E^2} = \frac{n\gamma^2 L^2 A_0^2}{R^2 C^2 \omega_0} \int_0^\infty \Phi(x) dx \tag{10}$$

This equation may be written in terms of the frequency, since $x = \omega/\omega_0$ and $\nu = \omega/2\pi$. We thus obtain:

$$\overline{E^2} = \frac{2n\gamma^2 L^2 A_0^2}{R^2 C^2 \omega_0^2} \int_0^\infty \Psi(\nu) d\nu. \tag{11}$$

But $\int_0^\infty \Psi(\nu) d\nu$ is the area under the curve obtained by plotting the relative amplifier out-put voltage, for constant input, against the frequency, and may be called A.

Thus:

$$\overline{E^2} = \frac{2n\gamma^2 L^2 A_0^2 A}{R^2 C^2 \omega_0^2}. \tag{12}$$

Circuit constants and the necessity of plotting the resonance curve may be eliminated by comparing this $\overline{E^2}$ with the normal shot effect $\overline{E^2}$ obtained from a comparison tube known to yield the accepted value of electronic charge, and in which the electron current has been adjusted to give the same $\overline{E^2}$ as the abnormal effect.

Thus:

$$\frac{2n\gamma^2 L^2 A_0^2 A}{R^2 C^2 \omega_0^2} = \frac{2 i_0 \epsilon L^2 A_0^2 A}{R^2 C^2}. \tag{13}$$ [10]

In the second member of this equation, ϵ is the electronic charge and i_0 the electron current necessary to yield the desired value of $\overline{E^2}$.

Simplifying, we obtain:

[10] N. H. Williams and H. B. Vincent, Phys. Rev. **28**, 1253 (1926).

$$\gamma = \omega_0 \left(\frac{i_0 \epsilon}{n} \right)^{1/2} \tag{14}$$

which is the expression for the average plate current increase due to the trapping of a single positive ion. From a consideration of the right hand side of the equation, it will be seen that we now have a means of determining. experimentally the size of the primary event of the phenomenon in question.

In the course of the experimental work it was found, as would be expected, that during the attack ofthe filament by oxygen the shape of the plate current characteristic of the tube was altered. The break-up of the space charge by the positive ions resulted in an increase in current in the range of plate potentials wherein the current was limited by space charge. The effect of this may be seen in Fig. 5. The method of taking this plate current characteristic will be discussed in a subsequent paragraph.

If we let Δi_P be the increase in plate current, at any one plate potential, due to the action of all of the ions, and let τ be the average time of life of an ion in the minimum of potential, we see that $\gamma\tau$ is the average charge released by one ion, and $n\gamma\tau$ is the charge released per second, which is precisely the total change in the current.

Therefore:

$$\Delta i_p = n\gamma\tau. \tag{15}$$

It may now be seen that, if n is known from the saturation value of the positive ion current with reversed field, Δi_P known by direct measurement, and γ computed from the shot effect determinations, substituted in Eq. (14), we have a means of computing τ, the average time of life of the ions in the minimum of potential.

In addition, it is of interest to determine a value of the number of electrons released by each positive ion. This may be had by dividing Δi_P by the saturation positive ion current under reversed fields.

It seemed possible that, from the value of τ, the magnitude of the space charge density in the minimum of potential might be computed and checked against magnitudes computed from a knowledge of the filament temperature and the space current. It was thought that this might offer a means of verifying our theory, in particular the applicability of the "flicker-effect" Fourier coefficients to this experimental case.

The writer is not aware of any exact formula connecting τ and the electron space charge density, but Kingdon and Charlton[11] have given the approximate relation:

$$\rho = \frac{\left(\frac{2\pi m_i}{KT_f} \right)^{1/2}}{4\pi b^2 \tau} \tag{16}$$

where m is the electron mass, T_f the filament temperature, and b is a parameter having the value 3.25×10^{-7} for a temperature of 1700°K.

[11] K. H. Kingdon and E. E. Charlton, Phys. Rev. **33**, 1012 (1929).

It is obvious that, in making use of this equation, we are assuming that the life of the positive ions is limited solely by recombination with electrons. The other principal factor which might limit the time of existence of the ions in the minimum of potential is escape from the ends of the electrodes. In order to test for this effect, a special tube was constructed, having shields over the ends of the plates which could be so charged as to repel the ions back into the region about the filament. It was found that the effect of loss by escape was not appreciable with the low gas pressures at which the latter work was carried out. The region of pressure in which it was necessary to work will be discussed in a following paragraph.

The method of computing the space charge density in the minimum of potential, from a knowledge of the filament temperature and the space current, is too well known to require description. We need only refer the reader to the papers of Langmuir[12] and Kingdon and Charlton.[13]

In the present work, we could not hope to check very closely the values computed by the last mentioned method, since Eq. (16) is not exact. However, since the rate of recombination must certainly vary directly with the space charge density, the variation:

$$\rho \sim \frac{1}{\tau} \tag{17}$$

may be expected to be true. In short, if our theory is correct, we would expect the values of ρ computed from the experimental results to behave, with some variable such as the space current, in the same manner as the values of ρ computed from purely theoretical considerations. If our theory were incorrect, this would certainly be most unlikely. This, then, yielded a means of testing the ideas outlined in this section.

Experimental Procedure. In the final experimental runs, made to test the foregoing theory, the oxygen pressure was maintained constant by pumping air through a slow-leak stop-cock and past the hot filament. Control experiments showed that the nitrogen and water vapor from the air were entirely inert and acted in no way to invalidate the work.

Determinations of the variation of the positive ion current with the oxygen pressure demonstrated a direct proportionality, from the lowest measureable pressures until the chemical action was so rapid as materially to decrease the surface of the filament during a run. At the higher gas pressures the number of positive ions formed per second was computed from direct measurement of the positive ion current under reversed field. At the low oxygen pressures, where the ion current was unmeasurable with a microammeter, readings of the current were taken at the higher pressures and the values extrapolated.

Before admitting oxygen to the tube, the vacuum plate current characteristic was taken. (See Fig. 5.) The utmost care was necessary in the determination of this curve, as the amount of oxygen given out by the

[12] I. Langmuir, Phys. Rev. **21,** 419 (1926).

[13] K. H. Kingdon and E. E. Charlton, Phys. Rev. **33,** 1011 (1929).

walls and electrodes, even after considerable baking and out-gassing, was sufficient to raise the characteristic curve in the region of space charge limited currents. This led to erroneous values of Δi_P.

Δi_P would be expected to vary linearly with the gas pressure, from Eq. (5), and a number of experiments were carried out to determine the range of gas pressures in which this was true. It was found that, with the tubes used, the linear variation held only up to pressures of the order of 10^{-4} mm of mercury. At higher gas pressures, Δi_P exhibited a saturation, no doubt partially due to the increased effect of the escape of the ions from the ends

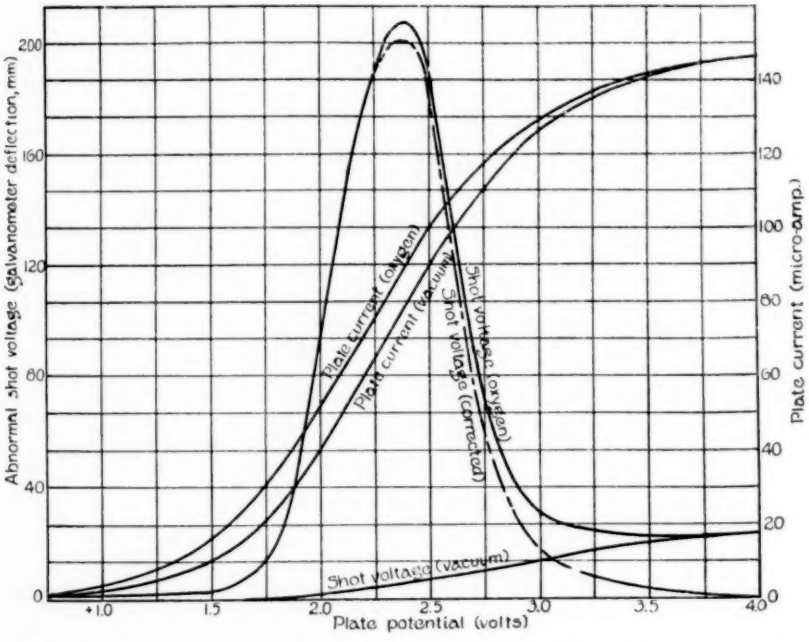


Fig. 5. Plate currents and mean square shot voltages as functions of the plate potential, in a vacuum and in oxygen.

of the electrodes. At the same pressure, ρ_m also began to fall away rapidly from the values computed by Langmuir's equations. For these reasons the pressures used in the final experiments were of the order of 10^{-4} mm of mercury or less.

Following the taking of the plate characteristic curve in a vacuum, the oxygen pressure was adjusted to the desired value and a determination of n made. The plate current characteristic of the tube was re-taken, together with the corresponding values of the abnormal mean square shot voltage. These latter were determined by means of the standard shot circuit and an amplifier which had been calibrated in terms of electron currents in a tube known to yield the accepted value of the electronic charge.

The values of γ, τ and ρ_m were computed from equations (14), (15) and (16), respectively. ρ_m was also computed from the equations of Lang-

muir and Kingdon and Charlton, and these latter space charge densities were compared with those obtained from the experimental results.

Results. Only one typical experimental run will be presented here. Fig. 5 illustrates the variation of the plate current in a vacuum and in oxygen, with the positive plate potential, and also the variation of the mean square shot voltage in both cases, with the same independent variable. The actual computations of γ, τ and ρ_m were carried out from the dotted curve.

TABLE II. *Values of η, γ, τ and ρ_m, computed from the experimental results illustrated in Fig. 5.*

E_p (volts)	η	γ (amperes)	τ (sec.)	ρ_m (experimental)	ρ_m (theoretical)
1.50	1.24×10^{5}	0.925×10^{-10}	2.13×10^{-4}	0.55×10^{9}	0.42×10^{9}
1.55	1.36	1.18	1.850	.637	
1.60	1.46	1.44	1.610	.728	
1.65	1.58	1.85	1.360	.861	
1.70	1.68	2.09	1.270	.920	
1.75	1.80	2.72	1.050	1.118	.790
1.80	1.94	3.27	.94	1.241	
1.85	2.06	3.94	.83	1.410	
1.90	2.22	4.92	.72	1.635	
2.00	2.38	6.98	.57	2.041	1.390
2.05	2.52	7.85	.52	2.230	
2.10	2.60	8.71	.45	2.630	
2.15	2.44	9.22	.41	2.820	
2.20	2.40	9.61	.38	3.080	
2.25	2.20	10.08	.35	3.360	2.130
2.30	2.14	10.32	.33	3.550	
2.35	2.10	10.52	.32	3.690	
2.40	2.06	10.31	.31	3.750	
2.45	2.00	10.08	.308	3.800	
2.50	1.80	10.00	.284	4.110	2.960
2.55	1.70	9.24	.293	3.990	
2.60	1.60	8.38	.303	3.860	
2.65	1.40	7.65	.291	4.02	
2.70	1.10	6.54	.267	4.38	
2.75	.90	5.65	.254	4.60	3.740
2.80	.80	4.95	.257	4.55	
2.85	.60	4.50	.212	5.50	
2.90	.50	3.96	.200	5.84	
2.95	.44	3.42	.205	5.70	
3.00	.40	2.92	.218	5.36	4.470
3.05	.34	2.83	.191	6.12	
3.10	.20	2.63	.121	9.65	
3.15	.16	2.42	.104	—	
3.20	.10	2.31	.069	—	
3.25	.04	2.06	.031	—	—

This last was obtained by first adjusting the vacuum shot voltage curve for the fact that the actual plate currents were higher in the second case, and then subtracting the result from the shot voltage curve obtained in oxygen.

The oxygen pressure chosen for this particular determination was 3.75×10^{-5} mm of mercury, which was well within the range of the linear variation of Δi_P with the gas pressure.

In Table II we have listed, for different plate potentials, the computed values of η (the average number of electrons released by each positive ion), γ, τ, and the experimentally determined values of ρ_m, which are the mag-

nitudes of the electron space charge density in the potential minimum, computed from the experimental results. In the last column are listed the values of ρ_m computed from the equations of Langmuir and Kingdon and Charlton. It will be noted that both η and γ reach a maximum, but not at the same plate potential, since η is proportional to the product of τ and γ. The values of τ are seen to decrease throughout the table, since the space charge (and hence also the amount of recombination with electrons) is decreasing. It will be noted that the values of τ are substantially greater than the period of the tuned shot circuit, which was in this case approximately 4×10^{-6} seconds.

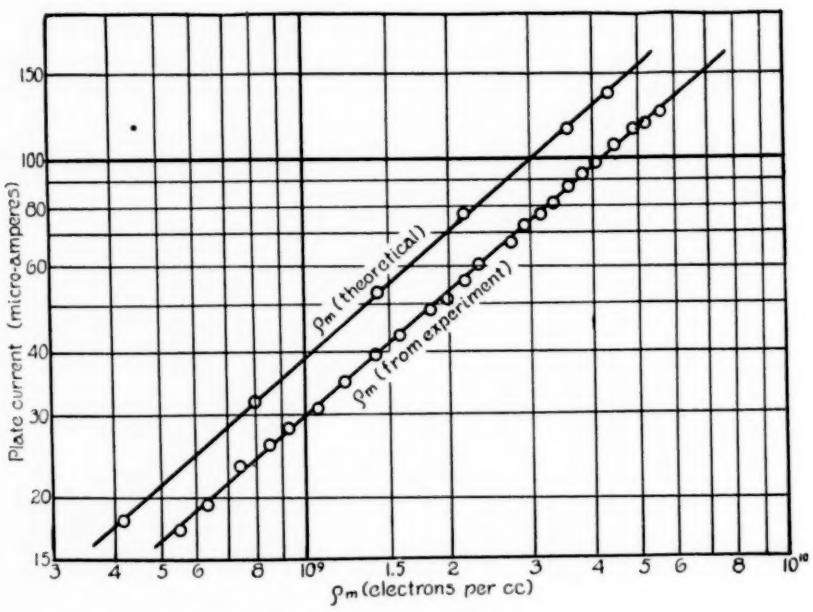


Fig. 6. Logarithmic plots of ρ_m, computed by the two different methods, against the logarithm of the space current.

In Fig. 6 we have the two different sets of values of ρ_m plotted against the space current, to logarithmic scales. As indicated in the last two columns of Table II, the values of ρ_m computed from the experimental results, agree with the values computed from purely theoretical considerations, within the probable error of the experiments and of Eq. (16). Fig. 6 shows that the two values of ρ_m varied in almost an identical manner with the space current, in spite of the fact that the experimentally determined magnitudes continue across the maxima of both the shot voltage and of Δi_P. (Fig. 5)

Summary and Conclusions

In the first part of this report, the formation of charged particles of tungstous and tungstic oxide was described. As far as the writer is aware, this has not been reported in the literature, and it is thought that the

results will be of interest to those working with the general problem of emission from various types of filaments.

In Part I we briefly described the abnormal shot effect resulting when the positive ions formed at the surface of the filament were accelerated through the dense electron space charge region surrounding the hot tungsten surface. This abnormality is interesting in that the fluctuations were not those of the direct current traversing the tuned shot circuit, since the large shot voltage was absolutely dependent upon the presence of the electron space charge. On the hypothesis that the primary event in the process is the fluctuation in the inter-electrode capacitance of the tube itself, due to variations in the thickness of the dense electron sheath, we were able to develop an expression for the mean square shot voltage in terms of the change in capacitance. Computations of ΔC from the experimental results yielded values which were entirely consistent with the actual electrode capacitance of the tube.

In Part II a study was made of the abnormal shot effect resulting from the trapping of the positive ions in the potential minimum surrounding the hot cathode. Tests of the frequency trend demonstrated that the phenomenon was of the nature of the "flicker-effect," the primary event persisting for a time longer than the period of the tuned shot circuit. From a consideration of the effect of the fluctuations upon the tuned circuit, an expression was developed for the experimental determination of the average increase in plate current caused by the trapping of a single ion. In addition, a means was found for determining the average time of life of an ion in the potential minimum, and the average number of electrons released by each ion.

As a means of checking the applicability of the theory, the values of the electron space charge density in the potential minimum were computed, and compared with the values obtained from a knowledge of the filament temperature and the space currents. The two sets of magnitudes of ρ_m were in agreement within the probable error of the experimental results and the equations used, and they were found to vary in exactly the same manner with the space current itself, throughout the range of plate potentials in which the currents were space charge limited. It is considered that these results lend strong support to the theory outlined in Part II.

As regards the technical aspects of the problem, it is obvious that the abnormal shot effects dealt with, and those mentioned in the introduction, are contributory factors to the noise in commercial vacuum tubes. The technique of abnormal shot effect should prove to be a powerful tool for the investigation of tube noise, and for the determination of the means of its elimination.

In conclusion, the writer wishes to express his indebtedness to Professor N. H. Williams, of this Department, under whose direction the work was carried out.

DISTINCTION BETWEEN CONTACT-POTENTIAL EFFECTS AND TRUE REFLECTION COEFFICIENTS FOR LOW-VELOCITY ELECTRONS

By H. E. Farnsworth and V. H. Goerke

Brown University, Providence, R. I.

(Received August 28, 1930)

Abstract

A method is described for determining the apparent electron reflection coefficient in two ways in the same apparatus such that a distinction may be made between a contact potential and a true reflection coefficient for very low-velocity electrons. The results indicate that for a copper target which has been heated at red heat for some time in a high vacuum, the true electron reflection coefficient approaches a zero value as the primary voltage approaches zero, and that the apparent increase in the reflection coefficient below about one volt, is due to a change in the contact potential between the target and the surrounding electrodes. The value of this change for the specimen tested is approximately 1.2 volts.

IN MAKING observations on total secondary electron emission from a metal target, it is found that after the target has been heated at red heat the *apparent* reflection coefficient* passes through a minimum at a bombarding potential of about one volt, and then rapidly increases as the voltage is further decreased.[1] In the past this has been attributed to a contact potential between the degassed target and the surrounding metal electrodes. Thus, if the target assumes a negative potential of one volt, say, with respect to its surroundings as a result of heat-treatment, then all primary electrons having an energy less than one equivalent volt are turned back from the target, and consequently the *apparent* reflection coefficient is increased. There is, however, the possibility that the effect in question is at least partly due to an appreciable *true* reflection coefficient for the degassed target at the voltage in question. From theoretical investigations by Nordheim,[2] in terms of the new quantum mechanics, it appeared that a large reflection coefficient may exist for very low-speed electrons, although a small but appreciable value was later predicted when taking into account the effect of an image force.

Former experiments have not permitted a distinction between an effect of contact potential and a *true* reflection coefficient. Most experimenters in this field have not concerned themselves with the voltage region as low

* The term "reflection coefficient" is used here in place of "coefficient of secondary emission" as in former papers by Farnsworth since all of the electrons leaving the target for the voltages in question are presumably reflected primary electrons. The word "*apparent*" is used to indicate that the reflection coefficient as observed by the method in question may not be the true one.

[1] H. E. Farnsworth, Phys. Rev. **20**, 358 (1922).

[2] L. Nordheim, Zeits. f. Physik **46**, 833 (1928); Proc. Roy. Soc. **A121**, 626 (1928).

as one volt because of the difficulty of obtaining a narrow beam of electrons of measurable intensity under these conditions. It, therefore, seemed desirable to attempt an experimental determination of the true nature of the above mentioned increase in the *apparent* reflection coefficient with decreasing voltage. Not only is this important from the theoretical aspect but the result enters into the determination of the experimental voltage of electron diffraction beams from single crystals, and hence into the evaluation of the inner potential of the crystal and consequently of the refractive index as previously emphasized by Farnsworth.[3] If the phenomenon is partly or wholly one of contact potential, then a correction must be applied to the impressed primary voltage, or the potential of the crystal must be changed by an amount equal to this correction. If the phenomenon involves only an actual reflection coefficient then no such correction should be made. Hence the accuracy of the experimental voltages is limited by the uncertainty of this correction.

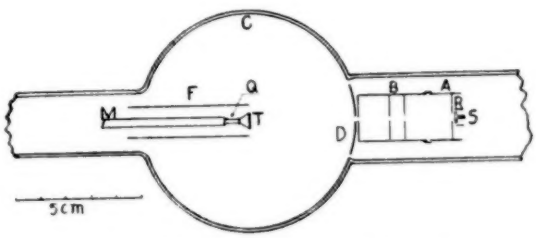


Fig. 1. Diagram of experimental tube.

Apparatus and Method

The essential parts of the apparatus are shown diagrammatically in Fig 1. The electron gun *SRAB* is similar to that previously described.[4] The molybdenum diaphragm *D* forms a part of the sphere *C*. The inner surface of the Pyrex sphere *C* is made conducting by evaporation from a molybdenum filament and contact is made by a tungsten wire, not shown, which is sealed through the sphere. The copper target *T* is supported by a quartz strip *Q* which is attached to a molybdenum frame *M*. The target and its support may be moved horizontally by means of a magnetic control into the side tube where the target may be heated at red-heat by high-frequency induction. The copper Faraday cylinder *F* may also be moved horizontally into the same side tube where it may be heated separately at red-heat by high-frequency induction. The target is a disk cut from the same piece of copper as the Faraday cylinder.

All of the glass parts are Pyrex with no ground joints or wax seals. The tube was kept attached through a liquid air trap to the vacuum system which consisted of two mercury diffusion pumps in series backed by an oil pump. The preliminary baking and vacuum conditions were similar to those in the previous experiments on secondary electron emission by Farnsworth.[4] No

[3] H. E. Farnsworth, Phys. Rev. **34,** 679 (1929).
[4] H. E. Farnsworth, J.O.S.A. & R.S.I. **15,** 290 (1927).

observable increase of pressure on a McLeod guage occurred during 24 hours while the pumps were not running. Further, the secondary electron characteristics of the copper target remained constant for more than 24 hours after heating at red-heat, thus indicating a negligible effect due to the residual gas.

The method of distinguishing between contact potential and a *true* reflection coefficient consists in observing the *apparent* reflection over the region of low voltages by the following two methods. (1) With the target T in the forward position, i.e. at the center of the sphere C, the secondary current from the target is the sum of the currents to C and D, while the total or primary current is that to all of the electrodes T, C, D, and F. Under these conditions only a very small current reaches the cylinder F by reflection from the sphere. The *apparent* reflection coefficient or ratio of secondary to primary current is thus obtained at small voltage intervals over the range in question. (2) With the target T removed to the back end of the cylinder F, the combination forms a Faraday absorbing cylinder. The total or primary current is the current to this Faraday cylinder. The current to the target T when in its forward position is obtained as in method (1). The difference between the total current and the current to the target is the secondary or reflected current, and the ratio of this to the total current is then the *apparent* reflection coefficient given by this method. It is seen that this *apparent* reflection coefficient will equal the *true* reflection coefficient if T and F are at the same potential.

The total current as obtained by this method should check with that given by method (1) if all of the metal electrodes T, C, D, and F are at the same potential. It will also check even if T is at a slightly different potential. In this case the electrons are all deflected to the inner surface of the cylinder instead of some striking the target. If, however, the cylinder F is one volt negative, say, with respect to C and D then no electrons having energies less than one equivalent volt will enter the cylinder so that the total current as obtained by method (2) will be zero while method (1) will indicate a total current greater than zero. Since the current to the target T in its forward position is the same in the two methods, it follows that the *apparent* reflection coefficients given by the two methods will not be the same. If, on the other hand, the *true* reflection coefficient of the target after red-heat treatment increases as the primary voltage is decreased below about one volt, and if no appreciable potential difference has been introduced by this heating, then the *apparent* reflection coefficients determined by methods (1) and (2) will be the same and equal to the *true* reflection coefficient for any given primary voltage.* A determination of the *apparent* reflection coefficient by these two methods for various primary voltages both before and after heating the target T and cylinder F at red-heat should then distinguish between a change

* This presupposes that all the electrodes T, C, D, and F were at the same potential before separately heating the target at red-heat. This was approximately the case since methods (1) and (2) gave the same result after baking the whole apparatus but previous to separately heating any of the electrodes.

due to an alteration in the contact potential and one due to an altered *true* reflection coefficient. This method depends for its success upon the elimination of a contact potential between the target T and cylinder F. For this reason these two electrodes were cut from the same piece of copper sheet and given heat treatments as nearly alike as possible.

The electron currents were measured by a sensitive galvanometer.

Results

After baking the whole tube but before heating either the target or cylinder at red-heat, measurements by methods (1) and (2) gave the same result. The reflection coefficient decreased with decreasing primary voltage, and appeared to approach a zero value for zero primary voltage. After heating

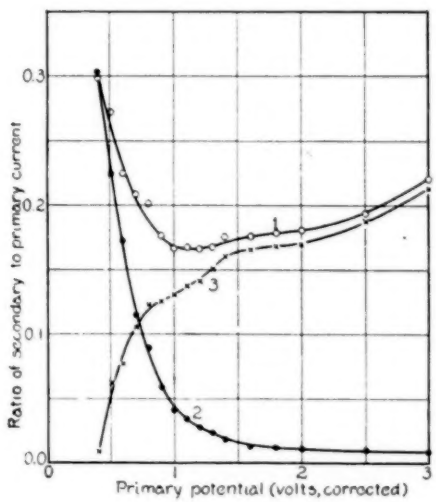


Fig. 2. Reflection coefficients by the two methods.

the target T alone at red-heat for several minutes, the formerly-observed increase in the *apparent* reflection coefficient, for primary voltages below about one volt, made its appearance. Both methods (1) and (2) gave the same result. This is shown in Fig 2 curve I for primary energies below 3 equivalent volts. After heating the cylinder F at red-heat, the results obtained by the two methods were not the same. With the target withdrawn to the back of the cylinder some of the low-velocity electrons were turned back from the cylinder and struck the sphere C and diaphragm D, thus resulting in an apparent reflection coefficient from the open end of the cylinder. The possibilities of the cylinder not being completely absorbing or of the spreading of the primary beam are eliminated by the results obtained previous to the heating of the cylinder for the primary velocities in question. Curve 2, Fig. 2 shows the *apparent* reflection coefficient from the open cylinder. Curve 3 shows the reflection coefficient from the target obtained by method (2). This was taken after the previous heat-treatment of the target and cylinder

were as nearly the same as possible. It is to be observed that this curve approaches a zero value as the primary voltage is decreased. The fact that it appears to cross the zero axis at a fraction of a volt probably has no significance. The voltages plotted represent the mean values which have been corrected for the velocity of electron emission from the filament, and for contact potential between the filament and other metal electrodes. Since the distribution of energies of the primary electrons covers at least one equivalent volt, the average value has little significance when only a few tenths of a volt are involved. Other curves corresponding to those shown in Fig. 2 were obtained when it was known that the conditions of the target and cylinder were not identical. In these cases the curves corresponding to curve 3 occupy positions between curves 1 and 2 of Fig. 2, as is to be expected. In all, the target and cylinder were each heated at red-heat temperatures such that excessive evaporation of copper occurred for periods totalling something over an hour. This heating removed sufficient gas to assure that the conditions of the target and cylinder were very nearly the same.

The results of this investigation indicate that for a copper target which has been heated at red-heat for some time in a high vacuum, *the electron reflection coefficient approaches a zero value as the primary voltage approaches zero* and that the *apparent* increase in the reflection coefficient below about one volt, which is observed only after red-heat treatment of the target, is due to a *change in the contact potential* between the target and the surrounding electrodes. The value of this change, which depends on the previous heat-treatment, is approximately 1.2 volts for the copper specimen tested.

CURRENT, PRESSURE, AND FREQUENCY RELATIONSHIPS FOR THE INITIATION AND MAINTENANCE OF THE ELECTRODELESS GLOW DISCHARGE

By Milton L. Braun

University of North Carolina

(Received July 3, 1930)

Abstract

Critical current and critical pressure values for the initiation of a discharge and their relation to the excitation frequency have been investigated by means of the electrodeless glow discharge for hydrogen, oxygen, carbon dioxide, and nitrogen. A departure from the usual definition of the critical point has been made in choosing as the criterion for its location the approximate center of the bend on the graph showing the relation between the current necessary to start the discharge and the pressure. By this method it was found: (1) that the critical current varies as the square of the wave-length; (2) that the critical pressure varies as the square of the wave-length; and (3) that the ratio of the critical current to the critical pressure is constant for a given gas. The current necessary to maintain the glow was also investigated, and formulas obtained expressing relationships between maintaining current, initiating current, wave-length, and pressure.

Introduction

RELATIONSHIPS between the minimum potential necessary to initiate a luminous discharge of electricity and the pressure of the gas through which the discharge takes place have long been investigated for several types of discharge. The pioneer work of Paschen[1] with spark discharge between parallel plate electrodes resulted in the generalization known as Paschen's law; namely, that for all values of pressure down to a limiting critical value, the minimum potential necessary to create a discharge is proportional to the product of the distance between the electrodes and the pressure of the gas. More recent investigations have shown that there exists a certain pressure for which the potential required to initiate the discharge is a minimum. This condition is referred to as the critical point.

The existence of critical points may be inferred from the works of several experimenters who had objectives other than the determination of critical values. Among such studies may be mentioned that of Frey,[2] who used a high potential battery discharge between ordinary electrodes in various mixtures of hydrogen and nitrogen; that of Kirchner,[3] who used continuous-wave, high-frequency discharge in hydrogen, oxygen, neon, and air, with various distances between electrodes; and that of Huxley,[4] who used an undamped electrodeless discharge in the study of corona effects for helium and neon.

[1] F. Paschen, Ann. d. Physik **37,** 69–96 (1889).
[2] B. Frey, Ann. d. Physik **85,** 381–424 (1928).
[3] F. Kirchner, Ann. d. Physik **77,** 287–301 (1925).
[4] L. G. H. Huxley, Phil. Mag. **5,** 721–737 (1928).

Comparing the internal and the external electrode methods, with various distances between the electrodes, and using undamped oscillatory discharges in air, Gutton and his fellow workers[5,6] observed critical points for a number of excitation frequencies corresponding to certain wave-lengths between 12,000 m and 27 m (2.5×10^4 and 1.1×10^7 sec^{-1}), reporting in addition the existence of a critical point for a 6,000 km wave-length (a 50 cycle oscillation). However, they found no critical point for a 24.9 m wave-length (1.2×10^7 sec^{-1}) though extending the observation down to pressures as low as 0.005 mm. There seems to be no simple relationship between their critical points, except that in the case of internal electrodes both coordinates of the critical point gradually increase with an increase in the frequency of oscillation. C. and H. Gutton[7] found that with hydrogen instead of with air the existence of critical points was restricted to a decidedly narrower frequency range, those corresponding to 110 m and 20.5 m wave-lengths being respectively above and below the critical range.

Critical points for 320, 160, and 80 m wave-lengths were found by Hayman[8] while using external electrodes with helium excited by continuous-wave radiation, but none was observed by him for a 40 m wave-length. Townsend and Nethercot,[9] however, also working with a 40 m undamped excitation but with nitrogen instead of with helium, not only encountered a critical point but were able also to determine that the potential coordinate was about one third less for external than for internal electrodes, other factors, including the critical pressure (0.04 mm), being equal. The comprehensive work of Heidemann[10] with hydrogen showed that with a 450 m wave-length the critical potential was nearly twice as great with external than with internal electrodes, but that the corresponding critical pressure was about one third less. It was also shown by Heidemann that an increase in the frequency resulted in a decrease in the critical potential, while the change in the critical pressure was irregular, the trend being toward an increase.

That the damped wave electrodeless ring discharge also gives rise to critical points was shown by J. J. Thomson,[11] by Bergen Davis,[12] and by Mierdel.[13] The data given in their papers, however, are insufficient to warrant any comparative conclusions. Thomson states[14] that an increase in frequency results in an increase in the critical pressure. Although only two frequencies were used by Davis, his work showed that with the higher frequency both the critical potential and the critical pressure were increased. A continuous-wave electrodeless glow discharge in a cylindrical tube was

[5] C. Gutton, S. K. Mitra, and v. Ylostalo, Compt. Rend. **176,** 1871–1874 (1923).
[6] C. Gutton, Compt. Rend. **178,** 467–470 (1924).
[7] C. and H. Gutton, Compt. Rend. **186,** 303–305 (1928).
[8] R. L. Hayman, Phil. Mag. **7,** 586–596 (1929).
[9] J. S. Townsend and W. Nethercot, Phil. Mag. **7,** 600–616 (1929).
[10] E. Heidemann, Ann. d. Physik **85,** 649–686 (1928).
[11] J. J. Thomson, Phil. Mag. **4,** 1128–1160 (1927).
[12] B. Davis, Phys. Rev. **17,** 501–505 (1903); Phys. Rev. **20,** 129–150 (1905).
[13] G. Mierdel, Ann. d. Physik **85,** 612–640 (1928).
[14] J. J. Thomson, reference 11, p. 1144.

recently used in a study of critical points for neon, helium, and oxygen by Gordon and Dushman.[15] Although they employed only three wave-lengths their results taken as a whole, showed that, an increase in frequency reduced the values of the critical point, with the exception, in the case of oxygen, of the pressure component.

It should be noted that the critical point has usually been given in terms of pressure and of the potential difference between electrodes, either internal or external. In the case of the electrodeless discharge investigators either have employed some artificial method of obtaining a value of the potential, or have discarded its use, choosing in its place the value of the current in the circuit of the exciting coil. The latter method is employed on the reasonable assumption that the strength of the field set up is proportional to the current that sets it up, at least for constant frequencies. While admitting the desirability of having direct potential measurements it must be recognized that the reliability of measurement of the mean square current by hot wire instruments gives this method an advantage over methods which purport to indicate the potential difference between selected points or along chosen paths within the field, and which associate the discharge itself with these more or less arbitrarily chosen potentials.

In the present work the critical point was determined for several gases in terms of the pressure and of the mean square value of the current in a coplanar excitation coil, the electrodeless glow discharge taking place in a spherical bulb placed at the center of the coil. The minimum current to maintain the glow at given pressures was also investigated.

Apparatus and Procedure

The apparatus used in this experiment consisted of (1) a gas control system, and (2) an electric circuit producing continuous-wave, high-frequency radiation. Low pressures were obtained by means of a vacuum pump, and evaluated by means of a McLeod guage. The entire gas system (Fig. 1), including the discharge bulb of 6.5 cm radius with gas inlet at the bottom, was an integral one constructed of Pyrex glass. The inner surface of the bulb was partially freed from moisture and occluded gases by the application of heat, and while the bulb was still warm the desired gas, from steel tank storage, was slowly passed through the system under slight pressure. Then in order to impregnate the walls of the bulb with the gas under consideration the confined gas was subjected to a pressure of about two atmospheres. Within an hour the pressure was slowly released to the desired value by opening the upper stop-cock, with the pump running. During the taking of data for a given charge of gas the pump was in continuous operation, the stages of pressure being controlled by the stop-cock on the pump side of the gauge, the gas, of course, passing only in the direction of lower pressure, out of the bulb into the pump. It was found that the stop-cock constrictions retarded the complete equilization of pressure in the system for an appre-

[15] N. T. Gordon and S. Dushman, Phys. Rev. **33,** 632 (1929), abstract.

ciable time, and this was taken into consideration before the pressure was determined.

The field which excited the glow discharge was generated by a tuned-plate tuned-grid system (Fig. 2). Oscillations were set up by a UV-205-A transmitting Radiotron, the plate of which was fed by a 1000 volt generator, filament current being supplied by a center-tapped transformer. Current in

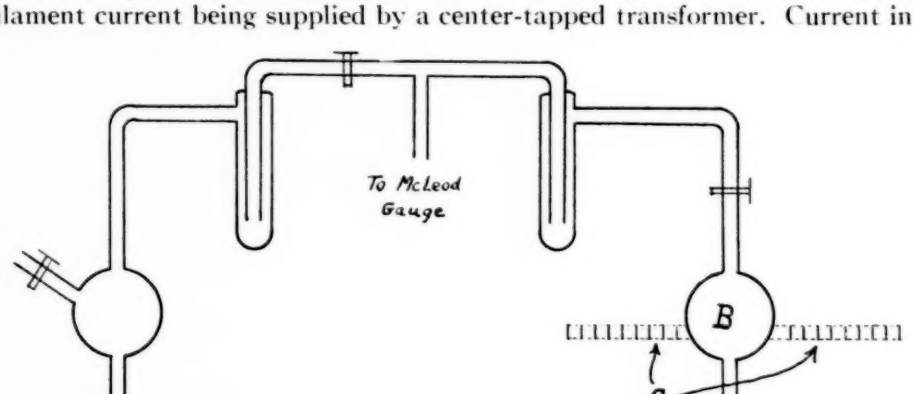


Fig. 1. Gas control system, showing radiating coil *C*, and discharge bulb, *B*.

the co-planar radiating coil was controlled not directly but by means of a suitable rheostat in the primary circuit of the filament transformer. Precaution was taken to keep the range of the filament current well within the linear section of the potential-current characteristic of the tube. The mean square value of the high frequency current was obtained by a hot wire ammeter having scale range from 0.2 to 2.0 amperes in .04 ampere divisions. The frequency

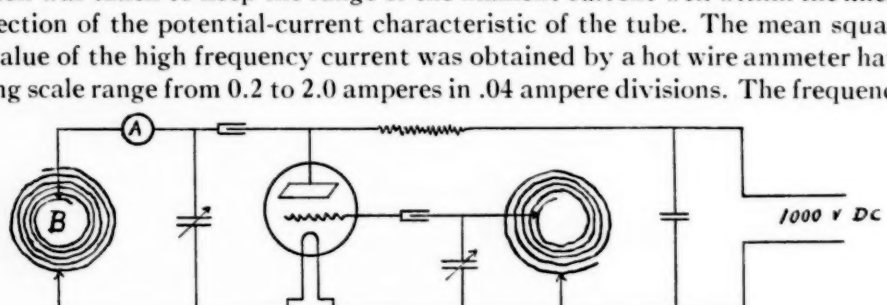


Fig. 2. Tuned-plate, tuned-grid, oscillatory circuit. Electrodeless glow discharge bulb placed at *B*.

of oscillation was controlled by manipulation of the condensers, while wave-length values were obtained directly from a precision wavemeter.

The initiating current was ascertained by adjusting the condensers for the predetermined wave-length, and, as the filament current was gradually increased, by noting its corresponding value in the radiating circuit at the moment when the visual discharge flashed on. The minimum maintaining

current was similarly obtained by noting, as the coil current was gradually diminished, its value just when the luminosity collapsed. Observations were taken within pressure limits of 2 to .002 mm of mercury in varying intervals which had their smallest values in the vicinity of the critical point. At each pressure, observations were made for wave-lengths between 90 and 160 m (or between 3.34×10^6 and 1.88×10^6 sec^{-1}) in five-meter steps. In order to assure simultaneous values in the variables, the pressure was read both before and after the several current observations were made. The pressure was then reduced, and another set of observations taken at each of the desired wave-

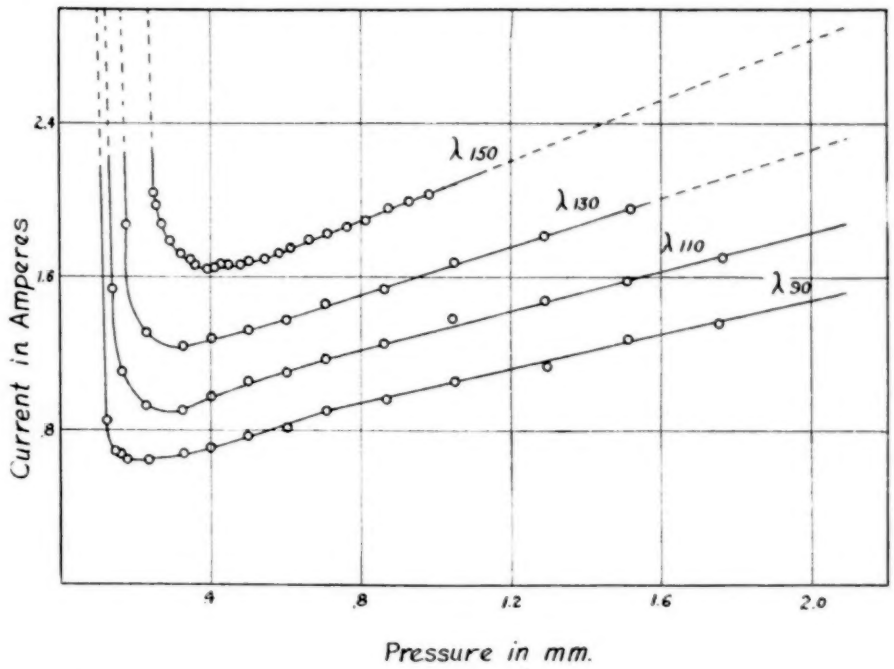


Fig. 3. Current to initiate the electrodeless glow discharge in nitrogen as a function of pressure.

lengths. Reliable results were obtained only after a time interval of 15 to 20 seconds from a previous discharge, it being found that an appreciable time was required for recombination in the gas to take place.

Minimum Current to Initiate the Glow

On plotting for a given wave-length the value of the initiating current as a function of pressure, it was found that for pressures greater than the critical, but not excessively greater, the current necessary to initiate the glow discharge is approximately a linear function of the pressure. This relationship is probably not a true linear one, for the curve tends to bend downward after a pressure several times the critical pressure is passed. Typical of the results obtained are the curves for nitrogen shown in Fig. 3. It will be seen that as the knee of the curve is approached from higher pressures, the regularity of the slope breaks down and passes into a rapidly changing one which seems

to rise asymptotically with the axis of ordinates. This indicates either that as the pressure is indefinitely decreased the initiating current is indefinitely increased, or that there is a pressure lower than which a discharge cannot be obtained. The knee of a representative curve is shown in detail in Fig. 4, which indicates values for oxygen on a 120 m excitation. For each gas and frequency investigated the curvature of the knee appears to be circular, and it is noted that the radius of curvature becomes smaller as the wave-length is decreased.

Determination of the Critical Point from the Data

Because of the rather large radius of curvature at the knee of the initiating current graph it was found impossible accurately to locate the critical point

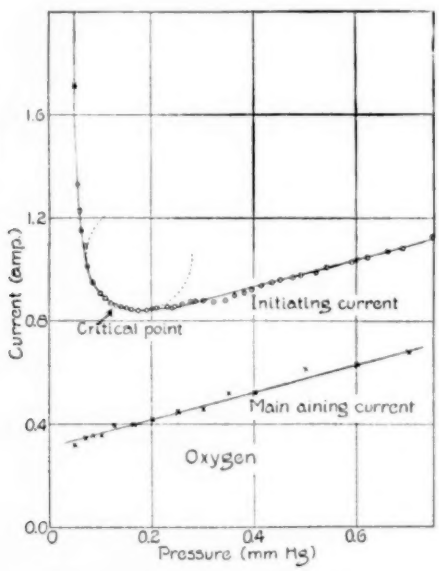


Fig. 4. Initiating current and maintaining current in oxygen as functions of pressure. Wave-length 120 meters.

by strict adherence to definition. This difficulty was also recognized by J. J. Thomson.[14,16] Critical points as arrived at by Davis[12] were determined by the intersection of lines tangent to either end of the knee, or by the intercept of the projected linear section of the curve on the axis of ordinates. In deciding upon the critical point the present work used for its criterion the central or bisecting point of the knee, thereby permitting its locus to remain on the curve.

The justification of this criterion for the location of the critical point lay in the simple relationships between the variables to which it led. In addition

[16] J. J. Thomson, "Conduction of Electricity Through Gases," Cambridge, 1st ed (1903), p. 362; 2nd ed (1906), p. 446.

to the difficulty of exact location, decided irregularities arose both between the points themselves and their relation to the frequency when a literal specification of minimum current and its corresponding pressure was followed. Values of the critical point as derived by this center-of-the-bend method are shown in Table I, where I_c is the critical current in amperes, P_c the critical pressure in millimeters of mercury, and λ the excitation wave-length in meters.

TABLE I. *Critical points for the initiation of the electrodeless glow discharge. Critical current, I_c, is measured in amperes; critical pressure, P_c, in mm of mercury; and excitation wave-length, λ, in meters.*

λ	Hydrogen		Oxygen		Carbon Dioxide		Nitrogen	
	I_c	P_c	I_c	P_c	I_c	P_c	I_c	P_c
90	0.44	0.110	0.50	0.070	0.63	0.132	0.68	0.155
95	.48	.120	.56	.086	.68	.140	.72	.168
100	.54	.135	.60	.094	.72	.150	.78	.180
105	.58	.150	.66	.100	.79	.165	.86	.200
110	.64	.160	.72	.105	.87	.185	.94	.215
115	.68	.170	.80	.120	.94	.200	1.03	.230
120	.74	.185	.86	.131	1.02	.215	1.09	.250
125	.80	.200	.92	.140	1.09	.230	1.16	.265
130	.86	.215	.99	.145	1.15.	.240	1.24	.280
135	.94	.235	1.08	.164	1.24	.260	1.33	.300
140	1.02	.255	1.15	.174	1.33	.278	1.42	.321
145	1.10	.275	1.25	.185	1.39	.290	1.47	.335
150	1.18	.295	1.34	.200	1.46	.305	1.62	.370
155	1.26	.315	1.40	.210	1.54	.320	1.68	.380
160	1.34	.340	1.48	.222	—	—	1.76	.400

CRITICAL POINT RELATIONSHIPS

The coordinates of the critical point being designated as the critical current and the critical pressure, it was found that for a given gas the critical current and the excitation wave-length may be expressed in terms of each other by the relationship,

$$I_c = a\lambda^2 + b, \tag{1}$$

where a and b are constants which have been evaluated for each gas, where I_c is the critical current in the oscillatory circuit measured in amperes, and where λ is the excitation wave-length in meters.

The relationship between the critical current and the wave-length does not in any way predict what the relationship between the critical pressure and the wave-length shall be. It was found, however, that for a given gas the critical pressure, too, is dependent upon the square of the wave-length in accord with the relationship,

$$P_c = c\lambda^2 + d, \tag{2}$$

where c and d are constants, P_c the critical pressure in millimeters of mercury, and λ the wave-length in meters. That the critical pressure increases with the square of the wave-length, or decreases with an increase of frequency, stands in contrast to the conclusions of Davis[12] and of Thomson[14] who found that for the electrodeless ring discharge an increase of frequency gave rise to an increase in the critical pressure.

Upon combining equations (1) and (2) it is seen that

$$I_c = kP_c + h, \tag{3}$$

where k and h are new constants. A graph of the critical points themselves not only gave a linear relationship, but also showed upon extrapolation that the straight line passed through the origin of the co-ordinate system. In other words it seems that within experimental limits the ratio I_c/P_c is the derived slope, k, itself. That the constant h does not come out to be zero is probably due to experimental error, and it is therefore concluded that

$$I_c/P_c = k \tag{4}$$

where k is equivalent to the ratio a/c from equations (1) and (2). Table II gives values obtained for the constants, while Fig. 5 shows the relationships between the three variables for one of the gases studied.

TABLE II. *Evaluated constants for equations* (1)—(4).

	Hydrogen	Oxygen	Carbon-Dioxide	Nitrogen
$a\times10^5$	5.18	5.72	6.00	6.39
b	0.012	0.037	0.145	0.166
$c\times10^5$	1.30	0.84	1.25	1.44
d	0.002	0.008	0.031	0.040
h	0.006	−0.016	−0.005	−0.011
$k=a/c$	3.98	6.82	4.80	4.43
$k=I_c/P_c$	3.99	6.67	4.77	4.39

MINIMUM CURRENT TO MAINTAIN THE GLOW

The minimum current necessary to maintain the glow for a given gas and frequency was found to be in direct proportion to the pressure. This relation showed no deviation at low pressures but continued linearly through the critical point to the low pressure limit of the apparatus. In order to obtain values of the minimum maintaining current for pressures lower than those for which the glow could be initiated, the discharge was effected at a higher pressure, and the pressure then reduced to the desired value while the discharge was in operation. The current was then gradually reduced until the glow became extinguished.

The minimum maintaining current was found to be a direct function not only of the pressure but also of the wave-length. The following empirical expressions give values for the minimum current, I_m amperes, required to maintain the glow, in terms of pressure and of wave-length:

$$\left.\begin{aligned} \text{Hydrogen } I_m &= (0.0044 + 0.0023P)\lambda - 0.20 \\ \text{Oxygen } I_m &= (0.0054 + 0.0044P)\lambda - 0.34 \\ \text{Nitrogen } I_m &= (0.0057 + 0.0038P)\lambda - 0.40 + 0.16P \end{aligned}\right\} \tag{5}$$

The maintaining current for the carbon dioxide glow was not investigated.

Perhaps of more importance than the relations determining the maintaining current are the connecting relations between the maintaining and the initiating currents. In presenting these relationships it is noted first that, for a given wave-length, and for pressures which lie within the restricted linear

section of the initiating current curve, there is a constant difference between the maintaining current and the initiating current. This difference may be expressed as

$$I_m = I_i - w \tag{6}$$

where I_m and I_i are the minimum maintaining and initiating currents respectively, and where w is the difference between them. The value of w was found to depend upon the nature of the gas, and the wave-length. The smallest observed value of w occurred with hydrogen at the shortest wave-length, while the greatest difference between the currents was found in the case of nitrogen at its longest wave-length. A representative relationship

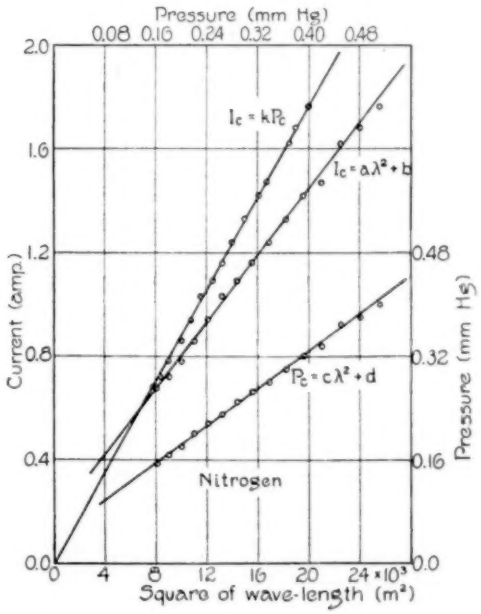


Fig. 5. Typical relation between critical current, critical pressure, and wave-length.

between maintaining and initiating currents is shown in Fig. 4. Including wave-length as one of the variables, the relationship between the two currents for the three gases investigated was found to be as follows:

$$\left.\begin{aligned} \text{Hydrogen } I_m &= I_i - 0.0044\lambda + 0.20 \\ \text{Oxygen } I_m &= I_i - 0.0062\lambda + 0.26 \\ \text{Nitrogen } I_m &= I_i - 0.0064\lambda + 0.19 \end{aligned}\right\} \tag{7}$$

The permissible values of I_i, of course, are those for which the linear relationship between I_i and P is best satisfied.

It is beyond the scope of this article to propose a theoretical explanation of the observed phenomena or of the relationships deduced therefrom. In conclusion the writer expresses his sincere thanks to Dr. Otto Stuhlman Jr. for suggestions given.

THE ABSORPTION OF SLOW HYDROGEN POSITIVE RAYS IN HYDROGEN

By Robert E. Holzer
University of California, Berkeley

(Received August 26, 1930)

Abstract

An experimental method for studying the absorption of slow hydrogen positive rays in hydrogen is described. The absorption coefficients of H^+, H_2^+ and H_3^+ in hydrogen were measured in the region between 60 and 850 volts. The absorption coefficient of H_3^+ was found to be smaller than that of H_2^+ and very much nearer the absorption coefficient of H^+ in magnitude. The absorption coefficient of H_2^+ decreased from a value of 40 cm^2/cm^3 at 60 volts velocity to 20 cm^2/cm^3 at 850 volts and that of H_3^+ decreased from 17 cm^2/cm^3 at 60 volts to 12 cm^2/cm^3 at 500 volts. The absorption coefficient of H^+ remained nearly constant at 8 cm^2/cm^3, about one-half the kinetic theory value. No minimum of absorption was observed for any of the ions in the region investigated. Qualitative experiments upon the nature of the absorbing process indicate that the absorption of H_2^+ is probably due to neutralization while scattering is probably the most important factor in the absorption of H^+ and H_3^+ ions.

Introduction

THE first quantitative measurements on the absorption of very slow protons in hydrogen were made by Aich[1] in 1922. Using a Lenard type of apparatus he found that the collision radius of an H_2 molecule and a 25 volt proton was approximately the same as the kinetic theory value of the radius of the H_2 molecule. In 1925, Dempster,[2] using the type of apparatus which he had previously described, studied the absorption of 900 volt protons in hydrogen. Since the beam of protons was absorbed in a magnetic field, a proton which was neutral over part of its path would have the apparent radius of its path increased while one which lost velocity would have its apparent path radius decreased. Dempster observed no neutralization or loss of velocity equivalent to a change of 2 volts in the 900 volt accelerating potential as the pressure in the absorbing chamber was increased from 0.00017 to 0.008 mm of mercury. Dempster therefore concluded that at 900 volts the proton moved through several hydrogen molecules without apparent neutralization or loss of velocity, that between 900 and 25 volts there was a critical potential at which the proton lost its ability to penetrate the hydrogen molecule and that between 900 and 13 000 volts[3] there was another critical potential at which the proton acquired the ability to ionize the hydrogen molecule. He notes that "Electrons acquire the velocity of protons used in this experiment by falling through 0.5 volt and it is significant

[1] W. Aich, Zeits. f. Physik **9**, 372 (1922).
[2] A. J. Dempster, Proc. Nat. Acad. Sci. **11**, 552 (1925).
[3] E. Rüchardt, Ann. der Physik **73**, 228 (1924).

that electrons of this speed pass through many molecules without absorption."

While the work of G. P. Thomson[4] on protons of between 5,000 and 25,000 volts velocity does not lead to a direct measurement of the effective size of the hydrogen molecule, his conclusions are worthy of note. Measuring the scattering of a beam of protons in hydrogen by observing blackening of a photographic plate beyond the edge of the image of the geometrical beam, he found: (1) appreciable small angle scattering in the velocity range studied, (2) a rapid decrease of scattering between 5,000 and 25,000 volts, (3) single scattering below 0.01 mm, (4) proportionality between scattering and pressure between 0.001 mm and 0.012 mm and (5) an inverse cube law of force at scattering centers.

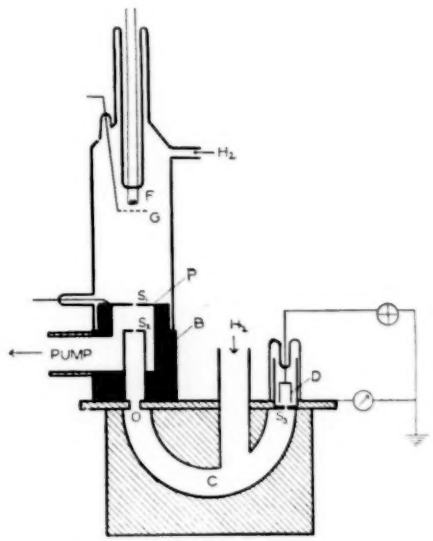


Fig. 1. Diagram of apparatus.

The present experiment was undertaken to measure quantitatively the absorption of protons in hydrogen in the region between the measurements of Aich and Dempster and to compare the absorption of protons with the previously unmeasured absorption of H_2^+ and H_3^+ in hydrogen.

Apparatus

The apparatus employed in the experiment consisted essentially in an ionization tube of the type used by Smyth,[5] Hogness and Lunn[6] and others for producing the hydrogen ions under controlled conditions and a Dempster[7] mass spectrograph for analyzing the rays. A simple diagram of the apparatus is shown in Fig. 1.

[4] G. P. Thomson, Phil. Mag. **1,** 961 (1926).
[5] H. D. Smyth, Phys. Rev. **25,** 452 (1925).
[6] T. R. Hogness and E. G. Lunn, Proc. Nat. Acad. Sci. **10,** 398 (1924).
[7] A. J. Dempster, Phys. Rev. **11,** 316 (1918).

Electrons from the incandescent filament, F, were accelerated toward the grid, G, by a potential difference of 60 volts. The primary ions were produced by electron impact in the space between G and the plate P. The spacing of F and G was 0.5 cm and of G and P, 10 cm. A drift field of 4 volts carried a small fraction of the ions formed between G and P through slit S_1, into the highly evacuated space between S_1 and S_2 where the major part of the accelerating potential was applied. The plate P was insulated from the adjacent metal parts B by a mica disk and picene wax. All metal parts in the ionization tube at ground potential (B) were completely covered by mica and wax.

S_1 and S_2, defining the beam of ions, were each 0.1 mm wide and 3 mm long. Their separation was 1 cm. The path of the beam between S_1 and O was shielded by a soft iron block, B, to which the ionization tube was waxed. The block, in turn, was soldered to the brass absorption and analyzing chamber placed between the pole faces of a large electromagnet so that the field was perpendicular to the plane of the diagram. When an ion passed S_2 it entered the absorption chamber and when it emerged from O it was deflected in a semicircular path the radius of which was determined by the mass and velocity of the ion and the strength of the magnetic field. If the magnetic field was so adjusted that the radius of the path was 4.75 cm the ion reached the slit S_3, 1.5 mm wide and 3 mm long, at the entrance to the collecting chamber. The total path length between S_2 and S_3 was 20.0 cm. A Faraday cylinder 0.5 cm in diameter and 1 cm deep was used to collect the ions. It was connected to a Compton electrometer and was insulated from the apparatus by 3 cm of Pyrex glass.

The pressure control was effected by steady flow through capillary leaks from reservoirs to the ionization and absorption chambers, respectively. The gas flowed from the chambers through S_1 and S_2 into the space evacuated by a Gaede three stage mercury vapor pump. Liquid air traps were included in all the vacuum lines and gas feeds to the apparatus. The pressures in the ionization and absorption chambers were practically independent. With a pressure of 0.0120 mm in the ionization, the "back pressure" built up in the absorption chamber was 0.0003 mm. A change of 0.006 mm in the absorption chamber produced a change of less than 1.5 percent (minimum detectable on the McLeod guage) in the ionization chamber. The pressure in the ionization chamber was approximately 0.01 mm and that in the absorption chamber, between 0.007 and 0.001 mm.

The gas was electrolytic (tank) hydrogen which was passed over hot copper turnings at 300°C, stored over P_2O_5 and finally passed through two liquid air traps after leaving the capillary while the pressure was less than 0.015 mm.

The magnetic field, calibrated by a test coil and Grassot fluxmeter, was found to be roughly proportional to the current flowing in the coil of the magnet when care was taken to reverse the magnet current before measurement.

The apparatus described produced a heterogeneous beam of hydrogen ions,

all of which had fallen through the same accelerating potential (i.e. within 4 volts). As the magnetic field was increased the rate of deflection of the electrometer reached several maxima corresponding to the various products of ionization. A typical graph of the results is shown in Fig. 2. All of these products of ionization have been observed previously.

Measurements

The absorption of the various ions was measured in terms of the absorption coefficient which appears in the well known absorption formula $I = I_0 e^{-\alpha x p}$, in which I_0 is the initial intensity of a beam of particles, I, the intensity of the beam after traversing a gas layer x cm thick at a pressure of

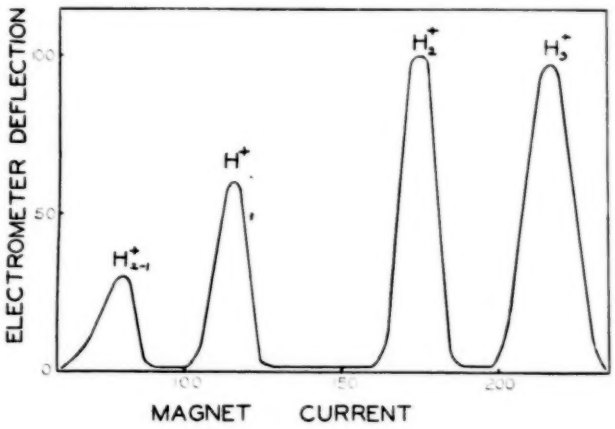


Fig. 2. Analysis of the products of ionization.

p mm, and α is the absorption coefficient, the effective absorbing cross-section of all of the molecules in 1 cubic centimeter of gas at 1 mm pressure. The equations

$$I_1 = I_0 e^{-\alpha x p_1} \tag{1}$$

$$I_2 = I_0 e^{-\alpha x p_2} \tag{2}$$

indicate the respective intensities of a beam originally of intensity I_0 after passing through a layer of gas, x cm thick, first at pressure p_1 and later at pressure p_2. Dividing (1) by (2) and taking the natural logarithm of both sides of the resulting equation,

$$\ln I_1/I_2 = \alpha x(p_2 - p_1).$$

Solving for α,

$$\alpha = \frac{1}{x(p_2 - p_1)} \ln \frac{I_1}{I_2}.$$

This equation is valid as long as I_0 remains unchanged and at once suggests a method of determining the absorption coefficient without measuring the absolute value of I_0.

In the experiment, I_0 was the intensity of a beam of ions of given mass and velocity that would have reached the Faraday cylinder had the absorption chamber been kept at high vacuum. I_0 was directly proportional to the intensity of the beam which entered S_2 and was roughly proportional to the current which reached the plate containing S_2. Since I_0 was a rather critical function of the filament current, the ionization tube pressure, and the applied potentials, keeping it constant was perhaps the greatest experimental difficulty encountered. A milliammeter measuring the differential current to P was used as an indicator of tube conditions while a galvanometer connected between the main metal portion of the apparatus and ground was used to determine the constancy of I_0.

The procedure in determining a value of α for a given ion at a particular velocity was briefly as follows. The tube was allowed to reach a steady state. The magnetic field was so adjusted that the maximum of one of the "peaks" (see Fig. 2) was observed. The rate of deflection of the electrometer (I_1) was recorded while the gas in the absorbing chamber was maintained at a given value p_1 (measured by a McLeod gauge). In practice the value of I_1 was taken as the average of three or four trials. The pressure was quickly raised by compressing the gas in the reservoir behind the leak to the absorbing chamber. Within a minute and half equilibrium was reached at the new pressure, p_2, which was 2 or 3×10^{-3} mm higher than p_1, and I_2 was measured. The pressure was then lowered to p_3, nearly the same as p_1, and the process was repeated. The set of measurements usually required from 7 to 10 minutes. The data thus obtained were sufficient to determine two values of α which served as an additional check on the constancy of conditions. In those cases where the values of α differed by more than 25 percent and the galvanometer showed signs of fluctuation, the measurements were discarded. If, however, the galvanometer deflection steadily increased or decreased, an average of the values of α was taken as a fair approximation to the true absorption coefficient.

It was necessary to be certain that the accelerating potential did not vary, for a change sufficient to shift a peak one quarter of its width was in most cases sufficient to render the coefficient meaningless.

The formula $I=I_0e^{-\alpha xp}$ holds when α is independent of the pressure or when $\log I \propto p$. Independent measurements were made to check this point. The relation was found to hold in the region studied, i.e., between 0.001 and 0.007 mm. This result is in accord with the results of G. P. Thomson for faster positive rays cited above. A further piece of evidence on this point is that α seemed to be essentially the same no matter what absolute values of p_1 and p_2 were used.

To obtain some idea of the processes involved in absorption of the ions a series of intensity measurements was made with a number of settings of the magnetic field for each of several pressures. The results of the measurements are shown graphically in Fig. 4. The meaning of the curves is briefly discussed in the following section.

Discussion of Results

The absorption of the various hydrogen ions in hydrogen is of interest since H^+ resembles the electron more than any other positive ion, H_2^+ is similar to H_2 in structure and H_3^+ has a mass comparable to that of H_2 but is unlike it in structure.

The absorption coefficients of H^+, H_2^+, and H_3^+ in hydrogen were observed in a region between 60 and 850 volts. The actual velocity ranges studied were 1.1 to 4.1×10^7 cm/sec for H^+, 0.76 to 2.9×10^7 cm/sec for H_2^+, and 0.62 to 1.79×10^7 cm/sec (corresponding to the range from 60 to 500 volts) for H_3^+. The results are shown graphically in Fig. 3.

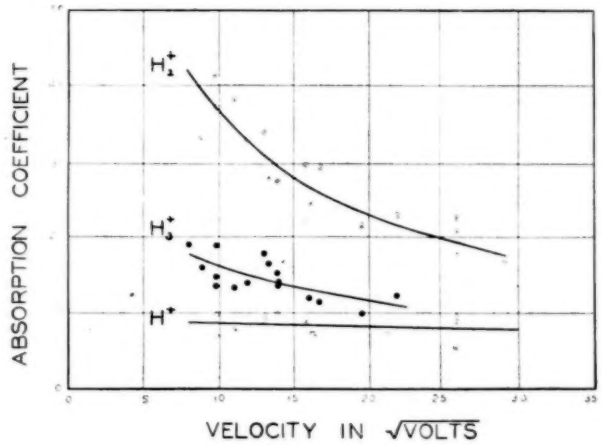


Fig. 3. Absorption coefficients.

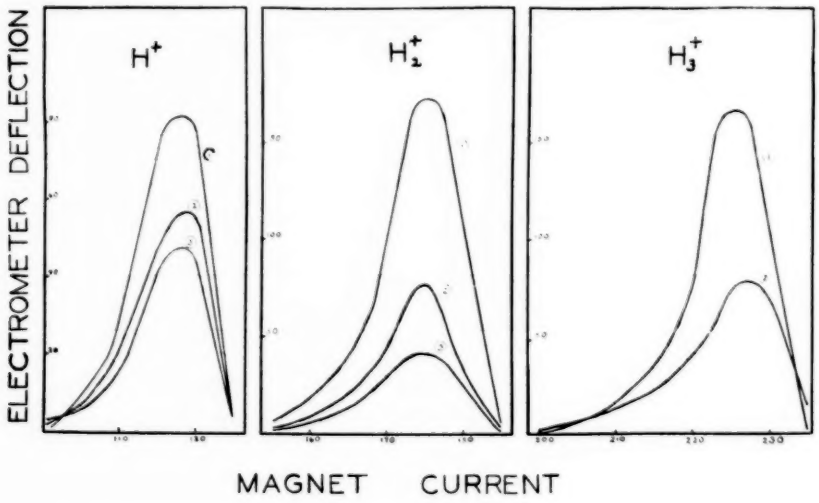
H_2^+ has markedly the largest coefficient varying from 40 cm^2/cm^3 at 60 volts to 20 cm^2/cm^3 at 850 volts. Assuming the radius of the ion given by Birge[8] and the kinetic theory value of the radius of the molecule, α would be 64 cm^2/cm^3. This value of α is definitely of the same order of magnitude as the experimental value of α at 60 volts. The absorption coefficient for H_3^+ is smaller and varies less, decreasing from 17 cm^2/cm^3 at 60 volts to 12 cm^2/cm^3 at 500 volts. The absorption coefficient for H^+ is nearly constant at 8 cm^2/cm^3, roughly half the kinetic theory value assuming H^+ of negligible dimensions.

In surveying the results, it is at once apparent that the kinetic theory "size" is not as important a factor in determining the absorption coefficient as the possibility of an energy transfer from the ion to the molecule. One would scarcely anticipate a lower value of α for H_3^+ in H_2 than H_2^+ in H_2 yet the value is not only smaller but far nearer H^+ than H_2^+ at the lower velocities. This result may be due to the fact that energy of ionization of an H_2 molecule is equivalent to the energy of neutralization of an H_2^+ ion and that a capture of an electron by an H_2^+ ion in H_2 is very probable. Since the

[8] R. T. Birge, Proc. Nat. Acad. Sci. **14**, 12 (1928).

same is probably not true of H_3^+, it is less likely to be eliminated by neutralization.

An examination of Fig. 4, showing the intensity and width of the H_2^+ peak at various pressures indicates that there is no spreading due to small angle scattering and no shift of the peak with increasing pressure. The process of absorption of the H_2^+ ion eliminates it completely from the beam. Since the free path is relatively long at the pressures used neutralization of the



Curve	H^+ $p\times10^3$	H_2^+ $p\times10^3$	H_2^+ $p\times10^3$
1	0.0	1.8	0.0
2	2.1	2.9	2.3
3	4.3	3.5	

Fig. 4. Relative absorption of the products of ionization.

ion would definitely eliminate it from the beam. Neutralization and some large angle scattering are evidently the most important factors in the absorption of H_2^+ in H_2.

The H_3^+ peak broadens with increasing pressure and diminishes in area. Since neutralization is not very probable the chief factor in the absorption of H_3^+ is presumably scattering. Some dissociation may occur although this point cannot readily be checked with the present type of apparatus.

The principal process in the absorption of H^+ is apparently scattering since in Fig. 4 the width at half-maximum does not decrease as the pressure increases and the bases of the peaks show a tendency to spread. Neutralization is not very probable since the energy of ionization of H_2 does not correspond to the energy of neutralization of H^+.

J. S. Thomson[9] has shown that α is a function of the geometry of the

[9] J. S. Thompson, Phys. Rev. **35,** 1196 (1930).

apparatus when small angle scattering is important. Since the results in Fig. 4 indicate very little small angle scattering, it is assumed that α is not markedly a function of the apparatus for any of the ions, less so for H_2^+ than either of the others.

The experimental value of α for H^+ at 60 volts is about one-half of Aich's value at 25 volts. These results are not discordant since one expects scattering to increase with decreasing velocities, very markedly in the region below which a minimum of absorption can be expected. However, the value of α does not show any sign of approaching zero at 900 volts as agreement with Dempster's results would require. Since Dempster gave no indication that the pressure in the ionization chamber was independent of the pressure in the absorption chamber, an increase in the intensity of the initial beam of hydrogen ions due to increased pressure in the ionization chamber may have masked the absorption effect.

The results of this experiment give no indication of a minimum of absorption of protons at 900 volts where the Handbuch der Physik [10] has interpreted Dempster's results as showing a Ramsauer effect. G. P. Thomson's[11] investigations of the scattering of protons in argon and helium indicate that the shape of the absorption curve for protons is similar to the shape of the absorption curve for electrons in these gases in the same velocity ranges. If this conclusion may be carried over to the case of the absorption of protons in hydrogen, a minimum of absorption is to be expected at 1800 volts instead of 900 volts, since Normand[12] has found a minimum of absorption for electrons in hydrogen at 1 volt. It is hoped that measurements of the absorption of 1800-volt protons in hydrogen can be made with the present apparatus in the very near future.

It is a pleasure to acknowledge my indebtedness to Professor R. B. Brode for suggesting this problem and for offering valuable advice and assistance throughout the course of the experiment.

[10] Handbuch der Physik, Bd. **24,** S. 100.
[11] G. P. Thomson, Phil. Mag. **2,** 1076 (1926).
[12] C. E. Normand, Phys. Rev. **35,** 1217 (1930).

THE VELOCITY DISTRIBUTION OF SECONDARY ELECTRONS FROM MOLYBDENUM

By Theodore Soller
Department of Physics, Amherst College

(Received August 25, 1930)

Abstract

The velocity distribution curves of secondary electrons from a molybdenum target for primary electron velocities from 7.5 to 100 volts are obtained by an improved apparatus, using the magnetic deflection method. Distribution curves quite similar to those observed by others for different metals are obtained, showing a broad low-velocity maximum of secondary electrons and a sharp "reflected" peak whose measured velocity in all cases turned out to be approximately 2 volts less than the corresponding accelerating voltage. It is shown that this should be so, from a consideration of contact potentials.

The effect of gas content of the surface on the velocity distribution was investigated. Heating at a temperature of approximately 850°C for 20 hours produced only slight changes in the curves. A prolonged heating of 41 hours at approximately 1250°C caused a marked decrease in the low-velocity maximum and an increase in the height of the "reflected" peaks. Both changes are ascribed to the removal of gas from the surface.

The curve giving the height of the "reflected" peaks as a function of the velocity of the primary electrons has several maxima. There is a fair correlation between the voltages to which these maxima correspond and the "critical potentials" of Mo observed by Petry.

Introduction

SEVERAL investigators[1–9] have been interested in the distribution in velocity of the secondary electrons which are emitted from metals when bombarded by a stream of electrons. The velocity distribution of these electrons has been most commonly obtained by subjecting the electrons in question to a uniform and known magnetic field of variable strength, and either measuring the number of electrons which reached a fixed collector as the field was varied, or else by receiving the electron spectrum on a photographic plate upon application of a constant magnetic field.

It seemed worth while to do further work along this line for the following reasons: (a) practically no attention has been paid to the effect of gas on the surface of the target upon the velocity distribution curves; (b) no work

[1] J. A. Becker, Phys. Rev. **23,** 664 (1924); **24,** 478 (1924).
[2] E. L. Rose, Phys. Rev. **25,** 883A (1925).
[3] A. Becker, Ann. d. Physik **78,** 228 (1925).
[4] C. F. Sharman, Proc. Camb. Phil. Soc. **23,** 523 (1927); **23,** 922 (1927).
[5] D. Brown and R. Whiddington, Nature **119,** 427 (1927).
[6] J. B. Brinsmade, Phys. Rev. **30,** 494 (1927).
[7] E. Rudberg, Proc. Roy. Soc. **A127,** 111 (1930).
[8] D. A. Wells, Phil. Mag. **5,** 367 (1928).
[9] H. E. Farnsworth, Phys. Rev. **25,** 41 (1925); **31,** 405 (1928).

of this nature has been reported on Mo, which lends itself readily to the process of outgassing because of its high melting point; (c) better vacuum conditions were desirable; and (d) the methods previously employed seemed to have one serious disadvantage in that no means of measuring the intensity of the incident beam of primary electrons was available. Consequently the curves obtained have been but apparent velocity distribution curves since nothing was known about the constancy of the primary electron beam impinging upon the target. The present experimental arrangement eliminates this difficulty.

Apparatus

Fig. 1 gives a schematic diagram of the apparatus and electrical connections. The metal parts of the apparatus were constructed entirely of molybdenum, with the exception of F_1 (Pt foil coated with oxides) and F_2 (tungsten). Electrons were accelerated between F_1 and E and then proceeded

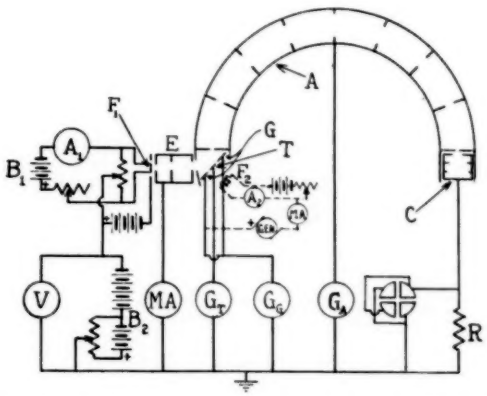


Fig. 1. Diagram of apparatus and electrical circuit.

to the target T, which consisted of two parts electrically insulated from one another, but in the same plane. Secondary electrons proceeding from the target in the direction of the first slit in the slit system A (called the analyzer) traveled in a circular path due to a uniform magnetic field produced by a pair of large Helmholtz coils and the earth's field. Those whose radius of curvature was 7 cm reached a collector C placed immediately behind the last slit. By varying the field, electrons of various velocities could be measured by the electrometer connected to C. Since the primary beam of electrons was also in this same magnetic field, their path would also be curved. Taking this into account, the angle between the direction of the incident electrons and that of the secondary electrons measured was in the neighborhood of 70–75°. The target was so constructed that only electrons from the central portion could reach the collector C. The target could be outgassed by electron bombardment from the filament F_2 behind it, without much danger of having the front surface contaminated by evaporated tungsten. Electron currents to G, T, and A were measured by galvanometers, as shown in the figure.

The metal parts of the tube were baked in a high vacuum before assembly, and great care was taken not to introduce contamination during assembly. The entire apparatus in its Pyrex tube was then attached to the pumping system and baked out at a temperature of 475°C for 30 hours. The pumps were kept running during the whole course of the experiment, pressures being read by a McLeod gauge sensitive to 10^{-6}mm Hg.

Method of Calculation

The velocity distribution was obtained by plotting the ratio of the current to the collector C to the current to the target T for various values of the magnetic field. As is well known, the deflections of the galvanometers G_G, G_T, and G_A do not measure the number of electrons striking the guard-target, target, and analyzer respectively, but rather the difference between this and the number that leave. The whole enclosure A, G, T may however be considered as being approximately a "black body" for electrons incident upon the target. Consequently, if we designate the currents as measured by the various galvanometers simply by the letters referring to those parts, then $(A+T+G)$ may be taken as a measure of the total primary current striking T and G. The number of electrons "sticking" to the whole target is measured by $(T+G)$. Hence the ratio $(T+G)/(T+G+A)=F$ gives the fraction of the total current striking T and G which is measured by the galvanometers G_T and G_G. In order to get the true electron current to the *effective* portion of the target, the current as measured by G_T, being T, must be multiplied by $1/F$. Hence the total effective primary current is

$$I_{target} = T\frac{T+G+A}{T+G}.$$

Furthermore, the ratio $I_{\text{electrometer}}/I_{\text{target}}$ for various values of the field, H, will not give the correct velocity distribution, because the effective width of the slit system varies directly with the velocity of the electrons whose velocity is to be measured. Hence, in plotting the curves, this ratio was reduced to the same effective slit width by dividing in each case by the value of the magnetic field. For convenience, this magnetic field was measured in tenths of gauss, and the ordinates of all curves are the ratio $I_{\text{electrom}}/I_{\text{target}}$ divided by the value of the magnetic field.

Method and Results

The velocity distribution curves shown in Fig. 2 were obtained after baking out the apparatus but without any special outgassing of the target. These curves are quite similar to those reported by Brinsmade[6] for Al. The essential features of these curves are a broad low-velocity group of electrons and a sharply defined high-velocity group. The value of the velocity of this group as calculated from the radius of curvature and the field strength in all cases turns out to be somewhat less than the accelerating potential as measured by the voltmeter V. That this should be so is evident

from the following considerations: The limiting case of a thermion from the filament is one which has just enough energy to escape from the filament. This may be designated as Φ_F, the work-function of the filament. This electron is then acted on by the field due to the applied e.m.f. V, by the contact e.m.f. $V_{C.TF}$ between the target and filament, and also by the field due to Φ_T, the work-function of the target, before it gets into the target. Therefore its kinetic energy in the target would be $e(V-V_{C.TF}+\Phi_T)$ $=e(V-\Phi_T+\Phi_F+\Phi_T)$ since $V_{C.TF}=\Phi_T-\Phi_F$, and consequently the kinetic energy in the target is $e(V+\Phi_F)$. The secondary electrons measured would not have a kinetic energy equal to this, however, since they would lose kinetic energy equal to $e\Phi_T$ in emerging from the target, neglecting any loss of energy while in the metal. Hence the kinetic energy measured would be given by $e(V+\Phi_F-\Phi_T)=e(V-\{\Phi_T-\Phi_F\})=e(V-V_{C.TF})$ and one would not expect

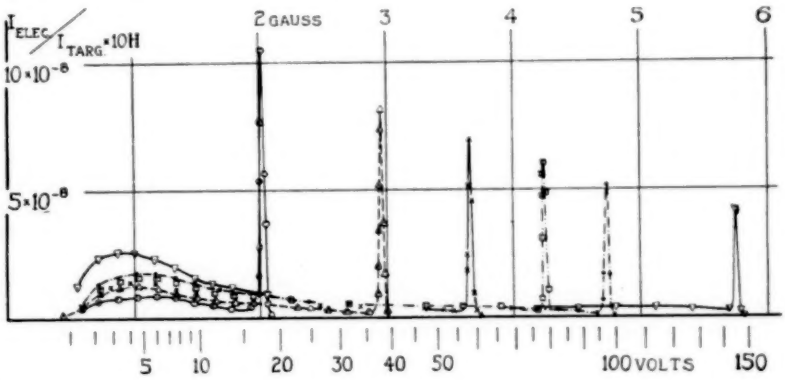


Fig. 2. Curves showing the velocity distribution of secondary electrons from a Mo target after baking out apparatus at 475°C for 30 hours.

the equivalent velocity (in volts) of the swiftest secondary electrons to coincide with V, the applied voltage, but to differ from this by an amount $\Phi_T-\Phi_F$ or $V_{C.TF}$. The above assumes no contact potential between the target and the analyzer. If there be a contact potential between T and A, $V_{C.AT}$ $=\Phi_A-\Phi_T$ then the kinetic energy of the electrons after entering the analyzer slit would be $e(V-V_{C.TF}-V_{C.AT})$ or $e(V+\Phi_F-\Phi_A)=e(V-\{\Phi_A-\Phi_F\})$ $=e(V-V_{C.AF})$. This contact potential $V_{C.AF}$ between the analyzer and filament is independent of changes in Φ_T due to the outgassing or other causes. In the present work, the accelerating voltage was very nearly 2 volts higher than the measured velocity of the electrons comprising the high-velocity peak. This would indicate that Φ_A is greater than Φ_F by 2 volts, which is a reasonable order of magnitude, and in the right sense. It would be reasonable to suppose, then, that the electrons which make up this peak are those which correspond in energy to the primary electrons.

The low-velocity group of electrons would include, as Sharman[4] suggests, electrons released by any characteristic corpuscular radiation excited in the atoms of the target by the primary beam, and also products of collision between primary electrons and free electrons in the target, as well as with

gas on the surface. It will be noticed that the magnitude of the low-velocity group increases with increasing velocity of the primary electrons, and there also appears to be a shift in the maximum toward higher velocities at the same time. These observations agree with those of Brinsmade.[6] At the higher primary velocities there is a small region on the low-velocity side of the full-velocity peak in which very few electrons are found. This has also been observed by Sharman,[4] Brown and Whiddington,[5] and Brinsmade,[6] working with other metals. Recently E. Rudberg,[7] using an apparatus of very much

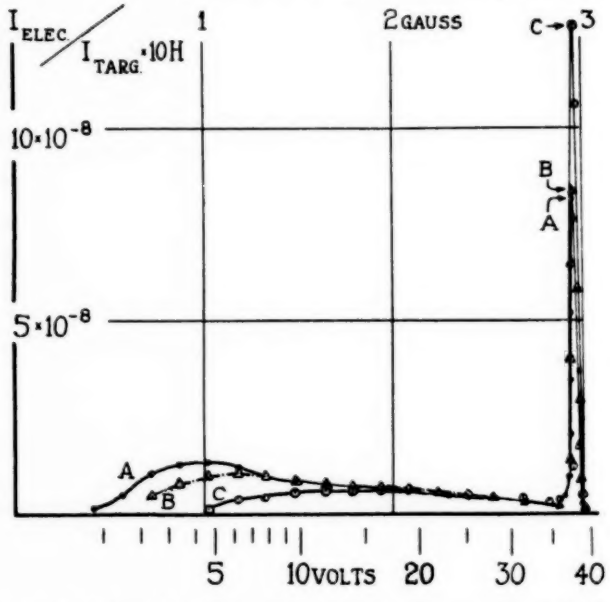


Fig. 3. Curves showing the effect of heat treatment on the velocity distribution of secondary electrons, for a primary accelerating voltage of 40 volts. *A*, after baking out at 475°C for 30 hours. *B*, after electron bombardment (850°) for 20 hours. *C*, after additional bombardment (1250°) for 41 hours.

higher resolving power and primary electrons of upwards of 200 volts, has shown that in this region there appear one or more maxima, indicating a certain kind of critical potentials for the substances studied.

The above runs were repeated after the target was outgassed for a period of 20 hours at a dull red heat. The differences between the curves thus obtained and the previous ones are not very great, the full-velocity maxima being practically the same in height, and the low-velocity group having its maximum slightly shifted toward higher velocities. The curve for 40 volts primary velocity is representative of what took place and is shown in curve *B* of Fig. 3, curve *A* being the original curve for this voltage. It is quite possible that heating at this low temperature does not cause a great change in the gas content of the surface itself, over what is produced by the original baking.

The target was then heated for a period of 41 hours at a much higher temperature, estimated at 1250°C, and during the last five hours of this the

temperature was raised still higher. During this last period the pressure in the system was too low to be readable on the McLeod gauge, even while the target was being bombarded. This condition was taken as an indication that this prolonged heat-treatment had removed practically all of the gas adsorbed on the surface of the target as well as a great share of the occluded gas. A series of curves was now taken as quickly as possible covering the range up to 100 volts at 5-volt intervals, in order to get more complete information as to the velocity distribution as a function of the primary velocity.

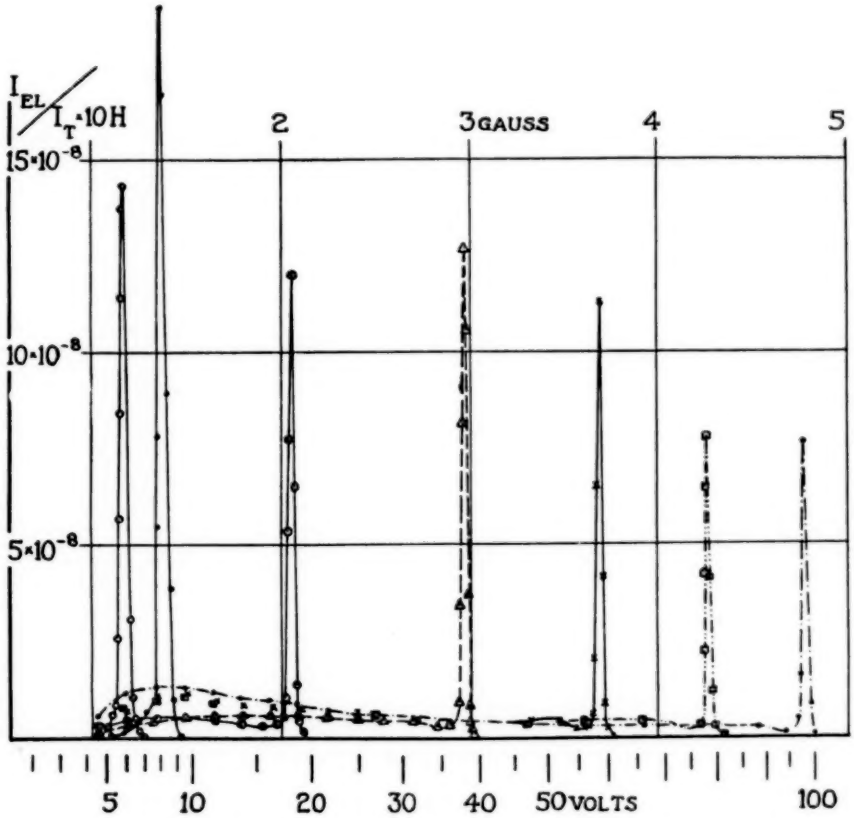


Fig. 4. Curves showing the velocity distribution of secondary electrons from Mo target after 61 hours heating at temperatures between 850° and 1250°C. The curves are for primary accelerating voltages of 7.5, 10, 20, 40, 60, 80 and 100 volts.

In the light of recent studies by Davisson and Germer[10] these curves cannot be interpreted as representing a gas-free surface condition, but the amount of gas accumulated on the surface when these measurements were taken must have been relatively small.

The changes introduced by this heat-treatment were identical for all velocities of primary electrons, and consequently it will suffice to give the results for the 40-volt curve. This is curve *C* of Fig. 3. It will be observed

[10] L. H. Germer, Bell Syst. Tech. Journ. **8**, 591 (1929).

that the full-velocity peak is considerably higher, and that the low-velocity peak has practically disappeared. Fig. 4 shows the curves for accelerating voltages of 7.5, 10, 20, 40, 60, 80, and 100 volts, after this last outgassing. The last five of these curves correspond in voltage to the first five in Fig. 2, and show clearly the changes both in the full-velocity peaks and in the low-velocity maxima. It will also be noted that for accelerating potentials under 15 volts practically all of the secondary electrons have the full velocity of the primary electrons. This would indicate either a pure reflection up to this point, or else perfectly elastic collisions. The separation into two groups begins to appear around 15 volts, and above this the separation is quite marked. This behaviour has also been noticed by Farnsworth[9] and Brinsmade[6] working with other metals.

The decrease in the low-velocity group upon heat-treatment can be explained if we suppose that the electrons comprising this group are the result of collisions between primary electrons and gas atoms. This group would then decrease as the number of gas atoms on the surface was reduced. The results of Brinsmade[6] also fit in with this hypothesis, since he obtained a maximum at almost exactly the same voltage from an Al target which had not been outgassed except by baking. If this maximum were due to a characteristic of the metal itself, one would not expect it to come at the same voltage for Al and Mo. The appearance of a slight low-velocity maximum for primary velocities above 40 volts may reasonably be due to collisions between primary electrons and atoms in the target, such that the resulting secondary electrons get through the surface of the target only after the energy of the primary electrons has reached a definite lower limit. This effect might reasonably be masked by the "gas" effect before thorough outgassing. The increase in the high-velocity or reflected group is also ascribed to the removal of gas from the surface, and parallels the experience of Davisson and Germer[10] who found that the reflected beams of electrons were much more pronounced when the target was relatively free of gas than when it was covered with many layers of gas atoms.

A rough idea of the relation between height of the full-velocity peak and the velocity of the electrons can be obtained by drawing a curve through the tips of the peaks of Fig. 4. In order to get better information regarding this point, the target was outgassed for 5.5 hours at a bright yellow heat, and immediately thereafter the readings were taken which are collected in the curves of Fig. 5. Curve 1 shows the first run, which extended to 40 volts, after which fluctuations in the battery voltage caused trouble. Curve 2 was obtained several hours after this, and extends to 100 volts. A third curve after still further outgassing failed to materialize because the last heating had evaporated enough tungsten to destroy the insulation between the target and its guard, making similar measurements impossible. The two curves of Fig. 5 agree pretty well in form, considering the difficulty encountered in trying to set exactly on the peak for each voltage.

One might expect that wave-interference maxima due to the microcrystalline structure of the target would show up under the conditions re-

presented in Fig. 5. Although some of the maxima can be accounted for in this way with reasonable assumptions, there are evidently more variables involved than were under control; consequently the details of computation are here omitted. Another interpretation of these maxima is that they may represent certain "critical potentials" of Mo. In comparing the positions of

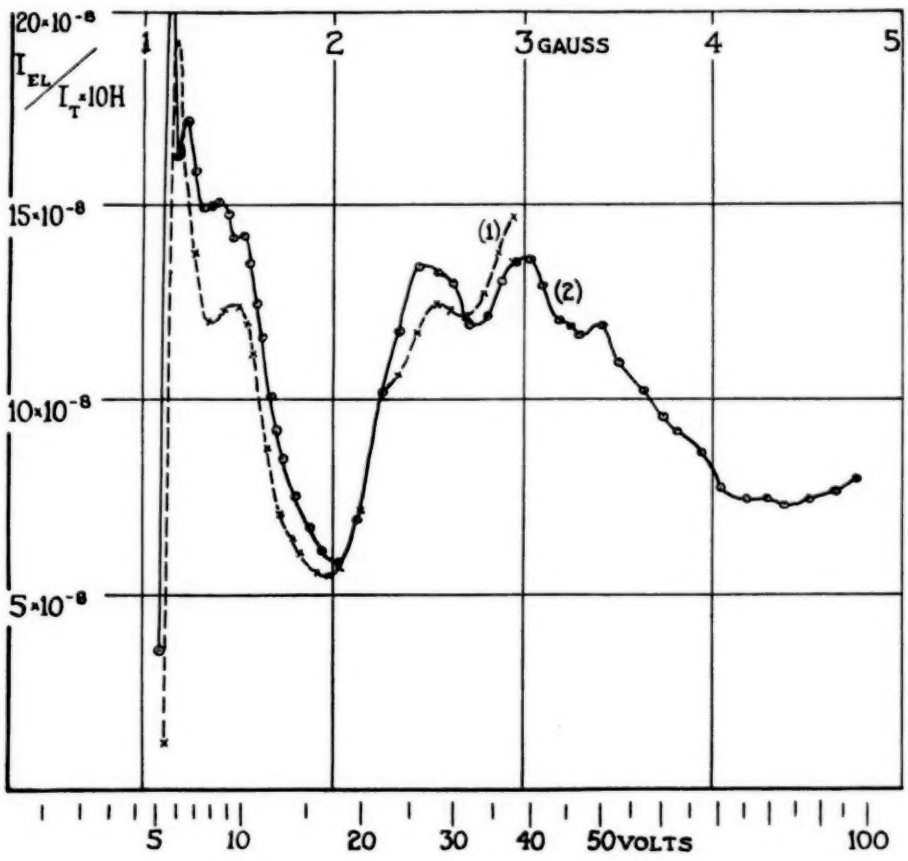


Fig. 5. Curves showing the heights of the full velocity peaks as a function of their velocity after thorough outgassing. Curve 1, first run. Curve 2, second run.

the maxima in Fig. 5 with the "critical potentials" which Petry[11] obtained in studying the ratio of total secondary to total primary electron currents for a Mo target, it is found that there is a very fair correlation between the two. Although the velocities at which the maxima occur do not coincide in the two cases, the velocities of the electrons in the target in both cases turn out to be very nearly the same, when proper corrections are applied to take care of the difference in work-function of the two sources, voltage drop along the filament, and initial velocity of emission of electrons.

[11] R. L. Petry, Phys. Rev. **26,** 346 (1925).
[12] J. M. Hyatt and H. A. Smith, Phys. Rev. **32,** 929 (1928).

J. M. Hyatt and H. A. Smith[12] working with a different type of apparatus, also find slight breaks in their curves for Mo, to which they attach no significance because they were not reproducible.

Further work needs to be done in order to prove or disprove this interpretation of the peaks. It does not seem probable that they can be due to eccentricities in the apparatus, since the irregularities appear only after considerable outgassing (Fig. 4) and not for the initial state (Fig. 2). The work is being continued at the University of Wisconsin, where the present work was done.

In conclusion, the writer wishes to express his thanks to Dr. C. E. Mendenhall for suggesting the problem and for his generous assistance throughout the work.

THE HEAT OF FORMATION OF MOLECULAR OXYGEN

By L. Covell Copeland[1]

Jefferson Physical Laboratory, Harvard University

(Received August 25, 1930)

Abstract

An apparatus has been constructed and described for the production of atomic oxygen and the direct determination of the heat of formation of molecular oxygen. A series of determinations have been conducted in the pressure range of 0.1 to 0.55 mm of mercury. The heat of formation of molecular oxygen was determined by these experiments as 131,000 (±6,000) calories or 5.7 (±0.3) volts per gram molecular weight. Several of the properties of atomic oxygen have been investigated including the possibility of long-lived metastable states.

Introduction

THE dissociation energy of molecular oxygen has been determined by several indirect methods, the probable values ranging from 110,000 calories or 4.8 volts to 162,000 calories or 7.05 volts. In this work a direct calorimetric determination has been made. In preliminary reports by Bichowsky and the author[2] the presence of atomic oxygen in the gas issuing from a discharge tube was first demonstrated and a few of its properties determined. The method employed for the determination of the heat of association of atomic oxygen is the same as that used by Bichowsky and the author in their determination of the heat of formation of molecular hydrogen.[3]

Electrolytic oxygen was admitted to the electrodeless discharge bulb D, Fig. 1, at a determined rate of flow by the capillary A. The gas partially dissociated by the discharge from the high frequency current in II, passes through a set of orifices B and is associated on the palladium black surface of the calorimeter E. The temperature rise of the calorimeter above the temperature of the constant temperature bath I gives the energy of association of the partially dissociated gas. The percent of atomic oxygen in the gas is determined from the formula $2\alpha/(1+\alpha)=3.41(1-P_N/P_D)$ where α is the percent dissociation, P_N is the pressure of the gas on the high pressure side of the orifices B under steady state conditions of no dissociation, before the discharge is turned on, P_D is the pressure of the gas at the same point under the steady state conditions involving dissociation when the discharge is going. As the development of this formula from Knudsen's formula for the rate of flow of gas through a small orifice is given completely in a former paper[3] it will not be repeated here.

[1] National Research Fellow in Chemistry.

[2] Bichowsky and Copeland, Nature **120**, 729 (1927); Phys. Rev. **31**, 1113(A) (1928).

[3] Bichowksy and Copeland, J.A.C.S. **50**, 1315 (1928).

Apparatus

The source of oxygen was a water-cooled electrolytic generator with nickel electrodes and CO_2-free potassium hydroxide solution as electrolyte. The rate of flow of gas into the apparatus was determined by capillary *A*. As the rate of flow of gas through a capillary is dependent on the pressure on the high pressure side, a manometer *M* was attached to the generator line which actuated a relay to control the D. C. current to the generator and thereby maintained the source of oxygen at constant pressure. Capillary *A* was also immersed in the constant temperature bath *I* (for simplification the sketch does not illustrate this). The rate of flow of oxygen into the

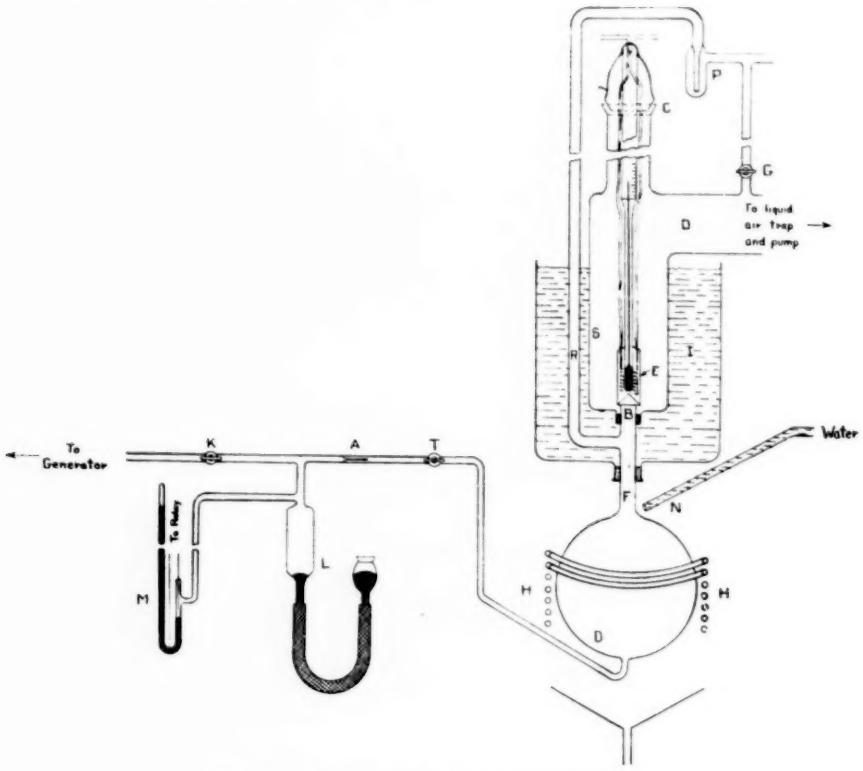


Fig. 1. Diagrammatic sketch of apparatus.

apparatus was measured by means of the pipette and leveling bulb *L*. The volume of the pipette between the two scratches on the capillary tubes on either end was accurately known. In determining the rate of flow, stop-cock *K* was closed and the time required to fill the pipette with mercury from the leveling bulb was taken on a stop watch, care being taken to maintain constant pressure as indicated by manometer and relay.

The discharge bulb *D* was made from a 500 cc round bottom flask. During an entire experiment this bulb was covered with running water from tap *N* discharging into the funnel placed beneath the bulb. The induction coil *H* was constructed of seven turns of 1/8 inch copper tubing

through which water was circulated to keep it cool. The current was supplied from an induction furnace circuit which consisted of 3kva 1–100 transformer, two 0.035 microfarad oil condensers and a quenched spark gap. The dissociated gas left the discharge bulb through the tube *F* which was a 1 cm piece of tubing about 13 cm long and terminated in the set of orifices *B*.

The use of Knudsen's formula for the flow of gas through orifices is applicable only if the area of the orifice is small in comparison to the mean free path of the gas, and the thickness of the plate in which the orifice is made must also be small in comparison to the diameter of the hole. Thus at 0.1 of a mm pressure the diameter of these orifices and the thickness of the wall in which they were made had to be of the order of magnitude of 0.1 mm.

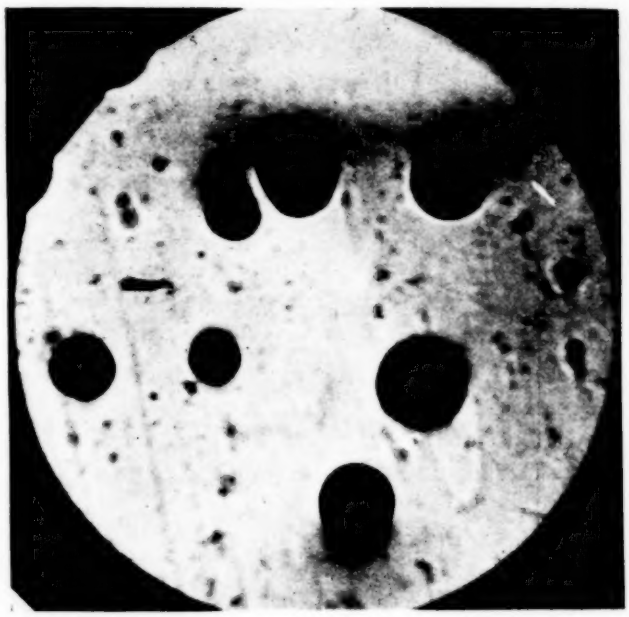


Fig. 2. Microphotograph of a portion of one set of orifices. Largest holes 0.2 mm in diameter.

Sets of orifices containing 15 to 40 such holes were obtained in the following manner. The end of a piece of Pyrex tubing was blown out to give a large fragile bubble of glass about 0.1 mm in thickness. A small section of this thin glass was placed on a carbon plate and the end of a piece of 1 cm cross-section Pyrex tubing was heated to flowing temperature and then pressed against the thin glass thus sealing it to the end. A piece of 1/32 inch tungsten wire was mounted in the chuck of a jeweler's lathe and a fine centered point produced by burning it with an oxygen flame while the lathe was rotating. This fine drill was then heated red hot with an air flame and the thin wall lightly brought to the hot point. After considerable practice several very satisfactory sets of orifices were obtained. A microphotograph of a portion of one of these sets is shown in Fig. 2 in which the largest hole is about 0.2 mm in

diameter. These orifices were sealed to tube F and fastened into the bottom of the calorimeter jacket S with sealing wax, to facilitate centering and replacement.

The pressure at B was measured on a large thermostated McLeod gauge which was connected to the apparatus through R, a tube of 1 cm diameter. This gauge was read with a cathetometer. A liquid air trap P was inserted in the McLeod gauge line in order to keep mercury out of the apparatus because it destroyed the atomic gas and poisoned the catalytic surface of the calorimeter.

Stop-cock G was closed during the run but opened as a bypass to protect the fragile orifices when it was desired to change the pressure in the apparatus by more than a few cms. This bypass was also of assistance when evacuating the entire apparatus before starting a run.

The calorimeter E was constructed of a cylinder of platinum foil 1.5 cm in diameter and 4.5 cm long. This cylinder was closed at the bottom end by a platinum foil cone and sealed to the bulb of a Beckmann thermometer with Wood's metal. A coil of enameled nichrome wire with enameled copper potential and current leads was sealed in the Wood's metal. This coil was for the calibration of the calorimeter. On the conical surface of the calorimeter was deposited electrolytically an even layer of palladium black. The Beckmann thermometer and calorimeter were sealed into the upper section of a ground joint C with sealing wax. Tube D served to conduct the gas to the conventional liquid air trap, mercury pump, fore pump, and low pressure McLeod gauge. The thermostat I consisted of an inverted bottle from which the bottom had been removed. It was fastened to F with a rubber stopper and supported separately. In addition to what is shown in Fig. 1, the thermostat contained a toluene thermoregulator, stirrer, lamp, Bechmann thermometer, cooling coil, and capillary A. The temperature of the water in the thermostat could be maintained constant to within 0.01°C.

Some Qualitative Properties of Atomic Oxygen

The first experiments[2] made on the study of atomic oxygen made use of a discharge tube containing aluminum electrodes for dissociating the gas. The electrodeless discharge produced a much larger percentage of atomic gas due to the absence of any metal that could produce a "clean up" effect. This fact is in accord with the observations of Kurt and Phipps.[4] Also in agreement with these authors it was found that water vapor was necessary for the establishment of dissociation. This fact was demonstrated very convincingly in the following manner. A 500 cc long-neck distilling flask was evacuated and then filled with moist oxygen gas to a pressure of about 0.2 mm and sealed off from the apparatus. When this bulb was held in the induction coil of the oscillating circuit a greenish blue discharge with a red center appeared that showed in the hand spectroscope, the line spectra of O I and the first three members of the Balmer series against a dark back-

[4] Kurt and Phipps, Phys. Rev. **34**, 1357 (1929).

ground. If the neck of the bulb was now placed in liquid air the entire character of the discharge slowly changed, the lines fading out and a banded background developing till finally $H\alpha$ the last to disappear, left nothing but a banded spectra. When the liquid air was removed, the reverse phenomena took place till at the end of a few minutes nothing but line spectra again remained. There were also a few specks of foreign material on the sides of the bulb that were incandescent when the line spectra were present and invisible when the band spectra were predominant. If the neck of the bulb was first cooled in liquid air and then placed in the exciting field, the first flash showed red and then immediately changed to the blue banded spectra. The interpretation is that the removal of water vapor by freezing it out with liquid air either prevented the dissociation of the O_2 molecule or, in analogy to Wood's[5] explanation of the role of water vapor in the production of atomic hydrogen, the removal of water vapor exposed the dry glass walls which were sufficiently catalytic for the recombination of oxygen atoms to cause their complete removal. In either case excited atoms were not present in the dry bulb in sufficient number to produce detectable line spectra or cause the foreign particles to glow by the heat of recombination of oxygen. The energy of the discharge was then taken up in the excitation of the molecules as shown by the ensuing band spectra. As another illustration of the necessity of water vapor it should be mentioned that when liquid air was applied to the inlet line of the apparatus between *A* and *D* during the discharge there was a gradual decrease in pressure and a steady drop in the temperature of the calorimeter indicating a decreased percentage of dissociation. Both of these readings returned to their former value when the liquid air was removed.

Although these experiments do not show any optimum pressure for dissociation as mentioned by Kurt and Phipps,[4] it should be noted that any dependence of dissociation at the orifices on pressure is probably completely masked by the changes in the surface conditions of exit tube *F* which was frequently changed throughout the course of these experiments and by changes in the rate of flow of gas through the apparatus.

Mercury vapor is very readily oxidized to the yellow form of mercuric oxide by atomic oxygen. Unless liquid air is maintained on the traps on all lines leading to a source of mercury vapor the cooler portions of the apparatus soon become coated with a yellow sheen due to this oxide.

Traces of nitrogen give the familiar greenish afterglow due to the metastable molecules. At the same time some compound is formed, possibly an oxide of nitrogen which will travel through the liquid air traps and react with mercury leaving a dirty black deposit[6] that destroys the use of McLeod gauges. For the above reasons tank oxygen could not be used in these experiments.

[5] Wood, Proc. Roy. Soc. **102,** 1 (1922).

[6] This black deposit may be removed without disassembling the apparatus even if it can not be heated, by filling the contaminated section with a few tenths of a millimeter of hydrogen and passing a discharge through that portion with an induction coil such as is used in testing for leaks.

Atomic oxygen like atomic hydrogen will cause small specks of substances, catalytic for the recombination, to glow white-hot in the dissipation of the energy of this reaction. A rough qualitative study of five metals was made to find a suitable catalyst for the surface of the calorimeter. Wires of these metals about 0.002 inches in diameter and 1 inch long were mounted on magnetic probes in another apparatus so they could be moved towards or away from the discharge bulb in vacuum. It was found that palladium, platinum and nickel were readily heated to glowing by the atomic oxygen but that copper and tungsten* were not visible affected. Palladium was apparently the most active and both palladium and platinum wires could be melted by bringing them too close to the discharge or increasing the flow of gas suddenly. Although it was necessary to bring the wires to within 10 cm of the discharge bulb to start them glowing, after they were hot they could be removed as far away as 35 cm from the discharge bulb in the direction the gas was flowing where they would still glow visibly and apparently indefinitely. This initial lag which was always present if the wire for any reason became cooled is believed to be due to a surface layer of gas or water vapor which prevented the reaction taking place except at high concentrations of atomic gas. Once the reaction was started the temperature of the wire kept its surface free until it was allowed to cool again. Platinum was first chosen at the most practical catalyst for recombination but subsequent experiments showed it was not sufficiently effective. Every collision of an oxygen atom did not mean capture for recombination. (See blank experiments) Palladium black however both because of its greater area and its greater activity was found to be sufficiently effective.

Experiments on Heat of Formation

At the beginning of all experiments the capillary A was closed off by the stop-cock T and the bypass G was opened. The apparatus was then evacuated until the McLeod guage showed a "flat guage," approximately 10^{-6} mm. Stop-cock T was then opened and the bypass G closed. The following readings were taken periodically; the pressure on both sides of the orifices, temperature of the bath, the temperature of the calorimeter and the laboratory air temperature. Occasional checks on the rate of flow were made. When steady-state conditions were obtained, which generally required one to two hours, readings of all temperatures and pressures were taken every 10 minutes for a period from half an hour to an hour. If these readings showed no appreciable drift the discharge was turned on and the process of obtaining steady state readings repeated. After the discharge was turned off a new steady state set of readings was obtained and at the same time the calibration of the calorimeter was made by passing sufficient current from a 12 volt storage battery through a variable resistance and ammeter and the heating coil of the calorimeter to maintain the temperature of the Beckmann thermometer at the same point as it was during the discharge. These ammeter

* In atomic hydrogen tungsten wires glow and platinum wires do not glow except with a large excess of oxygen. (See Wood, ref. 5).

readings and the reading of the potential drop across the heating coil gave an amount of energy per minute which was equal to that received by the calorimeter per minute during the discharge. The following computations are for Experiment 1.

Rate of flow of oxygen corrected to normal temperature and pressure 0.66014 cc/min.

Average pressure at B during last half hour of steady state of discharge 0.3544 mm.

Average pressure at B during last half hour of steady state of no discharge 0.3099 mm.

$$2\alpha/(1+\alpha) = 3.41(0.3544 - 0.3099)/0.3544 = 0.4282$$

$$= 27.24 \text{ percent of dissociated } O_2$$

$0.660 \text{ cc/mm} \times 0.2724 = 0.180 \text{ cc dissociated } O_2/\text{min}$.

0.123 amps at 0.60 volts required to maintain the calorimeter at the temperature obtained during discharge.

$0.123 \text{ amps} \times 0.60 \text{ volts} \times 60 = 4.428 \text{ joules/min}$

$$= 1.058 \text{ calories/min}$$

$22{,}412 \text{ cc/gr mol. wt} \times 1.058 \text{ calories/mm} \div 0.180 \text{ cc/min}$

$$= 131{,}645 \text{ calories/gr molecular wt.}$$

Table I gives the recorded data and results of fifteen experiments and one blank experiment, which will be explained later, arranged in the chronological order in which they were performed. Column 2 gives cc/min which is the rate of flow of oxygen corrected to normal temperature and pressure. Column 3 (P_D) gives the average pressure of the steady state during discharge

TABLE I.

Experiment number	cc/min	P_D mm of Hg	P_N mm of Hg	Percent Dissociation	cal/min	ΔH cal/mol
1	0.660	0.3544	0.3099	27.24	1.058	131,645
2	0.660	0.3522	0.3109	24.94	0.964	131,267
3	0.660	0.3085	0.3087	0.00	0.000	
4	0.667	0.2036	0.1789	26.15	1.222	156,985
	0.667	0.2032	0.1789	25.67	1.200	156,987
5	0.667	0.2012	0.1792	22.92	1.048	153,611
6	0.378	0.1088	0.0998	16.46	0.454	163,244
7	0.373	0.1089	0.0995	17.35	0.474	164,201
8	0.815	0.2455	0.2189	22.70	1.272	154,045
9	0.111	0.3766	0.3245	30.86	0.203	132,886
	0.111	0.3757	0.3245	30.27	0.188	125,489
10	0.168	0.5555	0.4813	29.51	0.288	130,381
11	0.064	0.2152	0.1905	24.38	0.121	172,360
12	1.129	0.5533	0.4867	25.84	0.176	135,023
13	1.129	0.5560	0.4846	28.03	0.177	125,239
14	0.947	0.4657	0.4150	22.76	0.130	135,197
15	0.947	0.4661	0.4163	22.06	0.127	134,633
16	0.361	0.1784	0.1626	17.87	0.456	159,382

and Column 4(P_N) the average pressure of the steady state after discharge. Column 5 gives the percent dissociation calculated from the pressure differences. Column 6 cal/min gives the energy in calories per minute required in the calibration of the calorimeter to bring it to the same tem-

perature as it was during the steady state of discharge. Lastly are listed the computed values of the heat of formation in calories per gram molecular weight. In Experiments 4 and 9 two separate steady states were obtained with the discharge on, involving different temperatures of the calorimeter. For this reason two separate computations were made on these experiments.

Blank experiments. In experiment 3 a strip of platinum foil coated with palladium black was inserted in tube *F*. The purpose of this foil was to associate all the atomic oxygen before it reached the orifices and calorimeter. The results of this experiment showed that there was no observable difference of pressure between the steady state of discharge and the steady state of no discharge and that there was no observable heat effect in the cal-

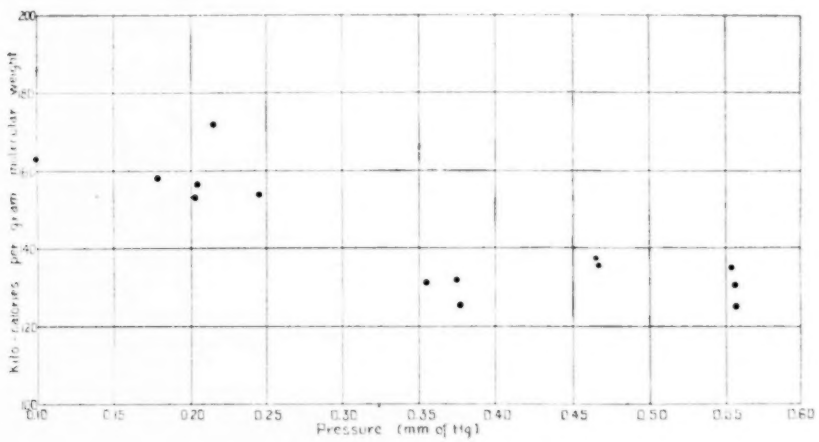


Fig. 3. ΔH-pressure plot of Table I.

orimeter. These results show that the observed pressure change in the other experiments was due to atomic oxygen which was completely associated by the palladium black surface to molecular oxygen and that there was no thermal leakage from the electrodeless discharge or exciting circuit to the calorimeter. These results were not obtained with a clean platinum surface in place of the palladium black. In the first experiments on this problem a shiny platinum surface was left on the calorimeter. Very inconsistent values were obtained for the heat of formation ranging from 50,000 to 150,000 calories. A similar blank with shiny platinum showed that only about 30% of the atoms were being associated. In all subsequent experiments the palladium black surface has been used. Whenever the calorimeter was removed a new surface of palladium black was deposited. This precaution was followed because one experiment not reported here, conducted after the apparatus had been standing for some time without liquid air on the traps, gave a value of 100,000 calories. Subsequent examination of the calorimeter surface showed an iridescently colored surface and minute drops of mercury visible with the aid of the microscope.

As a check on the possibility that the field of coil *H* might cause a temperature rise in the calorimeter not noticed in the blank experiment the oscillating current was allowed to pass through the coil when the pressure was too low to excite a discharge, 10^{-4} mm and also when the pressure was so high that only a faint glow was visible, 10 mm. In both cases there was no discernable temperature rise in the calorimeter. There was no appreciable stray discharge to the calorimeter. To demonstrate this fact an auxiliary electrode was sealed into a side arm on *F* and grounded through a galvanometer. No current passed through the galvanometer during the discharge.

In all these experiments the pressure of the low pressure side of the small holes was about 1/80th of that on the high pressure side.

Discussion

It should be noted that all values obtained from experiments where the pressure was below 0.3 of a mm are higher and not as consistent as the values obtained from those experiments conducted at higher pressures. Fig. 3 is a plot of all values of Table I in kilocalories as abscissas against pressure

Table II.

Author	Source	Value cal/gr mol.wt.
Warburg[8]	Photochemical formation of O_3	138,000
Born and Gerloch[9]	" " "	162,000
Eucken[10]	Heat of formation of CO	253,000–423,000
Wulf[11,12]	Ionization potentials	56,000–138,000
Hogness and Lunn[13]	Ionization potentials of O_2	150,000
Birge and Sponer[14]	Absorption bands of O_2	162,000
Birge[15,16,17,18]	" " "	138,000
Herzberg[19]	" " "	149,500
Kassel[20,21]	Decomposition of O_3	110,000
Mecke[22,23,24]	Predissociation in NO_2 bands	128,000
Henri[25]	Predissociation in SO_2 bands	126,400
Kaplan[26]	Birge's paper	131,000
Kondratjew[27]	Predissociation in NO_2 bands	118,000
Rodebush and Troxel[28]	Direct calorimeter measurement	131,000
Baxter[29]	Predissociation bands of NO_2	115,000
Average value of nine high pressure runs of this paper		131,300

[7] Copeland, J.A.C.S. **52,** 2580 (1930).
[8] Warburg, Zeits. f. Electrochemie **26,** 58 (1920).
[9] Born and Gerloch, Zeits. f. Physik **5,** 433 (1921).
[10] Eucken, Ann. d. Chemie **440,** 111 (1924).
[11] Wulf, J.A.C.S. **47,** 1944 (1925).
[12] Wulf, Proc. Nat. Acad. Sci. **14,** 615 (1928).
[13] Hogness and Lunn, Phys. Rev. **27,** 733 (1926).
[14] Birge and Sponer, Phys. Rev. **28,** 259 (1926).
[15] Birge, Phys. Rev. **27,** 641 (1926).
[16] Birge, Nature **122,** 842 (1928).
[17] Birge, Phys. Rev. **34,** 1062 (1929).
[18] Birge, Trans. Faraday Soc. (1929).
[19] Herzberg, Zeits. f. Phys. Chemie **4B,** 223 (1929).
[20] Kassel, Phys. Rev. **34,** 817 (1929).
[21] Kassel, Zeits. f. Phys. Chemie **2B,** 264 (1929).
[22] Mecke, Naturwissenchaften **51,** 996 (1929).

in millimeters of mercury as ordinates. In a former note[7] the value obtained at low pressures was presented. Since that time the last eleven experiments given here have been made. Since the high pressure values are consistent among themselves within the experimental limits of accuracy over a range of pressure from 0.35 mm of mercury to 0.56 mm of mercury, the average of these results is taken as the correct value. It is quite possible that the high values at low pressures are due to excess energy from metastable oxygen atoms which would have a longer life at lower pressures due to the fewer number of collisions. It was on this assumption that the high pressure measurements were made. Further confirmation of this point was sought by making a few runs with the discharge bulb moved 42 cm away from the calorimeter. Under these conditions it would be expected that there will be a decrease in the percentage of the atoms in the metastable state, by the time they reached the calorimeter. Also there will be a decrease in the total number of atoms. Unfortunately experiments made with longer paths have not proved to be sufficiently reliable to draw any conclusions.

Table II gives the reported values of the heat of formation of molecular oxygen and the experimental source from which the computation was made.

All values of experiments made at pressures above 0.3 mm of mercury agree within 6,000 calories of their average which is therefore taken as the experimental error.

It is of considerable interest that the preliminary report of Rodebush and Troxel is also from a direct calorimetric determination which differs but slightly from the method reported here. The agreement of our results is of more significance because of the differences in the method which include a fundamentally different application of Knudsen's formula for the measurement of the percent of dissociation. This result agrees also with that of Kaplan's and within the experimental limits with those of Mecke and Henri.

For a discussion of the theoretical assumptions involved in this method the reader is referred to the paper of Bichowsky and Copeland.[3] In addition to what is stated there concerning the validity of the use of Knudsen's formula it should be mentioned that in this work it was found that the ratio of the rate of flow of oxygen to the pressure was a constant for each set of orifices over the range of pressures used.

The author wishes to acknowledge his indebtedness to Dr. F. R. Bichowsky who first suggested this problem and the method used. He also wishes to acknowledge the valuable assistance received from Mr. Foster Rieke including the final perfection of the method of making the sets of orifices. He

[23] Mecke, Nature **125,** 526 (1930).
[24] Mecke, Zeits. f. Phys. Chemie **7B,** 108 (1930).
[25] Henri, Nature **125,** 275 (1930).
[26] Kaplan, Phys. Rev. **35,** 436 (1930).
[27] Kondratjew, Zeits. f. Phys. Chemie **7B,** 70 (1930).
[28] Rodebush and Troxel, J.A.C.S. **52,** 3467 (1930).
[29] Baxter, J.A.C.S. **52,** 3468 (1930).

is indebted to Professor L. C. Graton for the microphotographs. In conclusion he also wishes to thank Professor E. C. Kemble and the other members of the Physics Department of Harvard University for their valuable assistance in the direction of the activities of his fellowship.

THE DIELECTRIC CONSTANT OF HELIUM

By J. V. Atanasoff
Department of Physics, University of Wisconsin

(Received August 11, 1930)

Abstract

The most successful study of unperturbed helium in its normal state has been made by Hylleraas, using the Ritz method. The Schrödinger partial differential equation is reduced to an equivalent problem in the calculus of variations and this problem is then solved by a so-called direct method.

In the present paper an extension is developed to the Ritz method, forming a theory which is applicable to perturbation problems. This theory is applied to a study of the normal helium atom under the influence of an electric field. As in the usual perturbation theory, the characteristic energy parameter E and the characteristic function ψ are assumed developable in power series in the field strength F, but now the coefficients in these expansions are given in terms of problems in the calculus of variations. The work of Hylleraas furnishes a knowledge of the terms in both expansions independent of the field strength. Then the minimization of a single integral furnishes values of ψ_1 and E_2, the subscript of each denoting the power of F for which it forms the coefficient. Values of these coefficients are calculated to three approximations by the direct method. From E_2 we may calculate the dielectric constant in the usual way giving

$$\epsilon = 1.0000665,$$

a value 5 percent below the accepted value of $\epsilon = 1.000070$. One should note that it is a characteristic of the Ritz method that further approximations can only increase the value of ϵ, bringing it into better agreement with the experimental value.

THE value of the dielectric constant of an atom depends on the manner in which the normal energy levels are affected by a perturbing electric field. Unperturbed helium in the normal state has been most successfully studied, with the so-called Ritz method, by Kellner[1] and later by Hylleraas in two papers which we shall refer to as Hyl. I[2] and Hyl. II.[3] The work of Hylleraas in his second paper possesses the great advantages of simplicity and of rapidity of convergence. These advantages seem to be founded in his exclusive use of distances rather than angles as variables. In the present paper we shall employ his results for that description of unperturbed normal helium which forms the starting point for our perturbation calculation.

A. The Ritz Method and Perturbation Theory

§1. *The Ritz method and the characteristic value problem.* The ordinary way to solve a problem in the calculus of variations is to set up the corresponding Euler differential equations and find the solution which fits the boundary conditions. Under certain restrictions it is possible to reverse this procedure; i.e.,

[1] G. W. Kellner, Zeits. f. Physik **44,** 91, (1927).
[2] E. A. Hylleraas, Zeits f. Physik **48,** 469, (1928).
[3] E. A. Hylleraas, Zeits. f. Physik. **54,** (5, 6) 347 (1929).

given a differential equation, it is possible to find the corresponding problem in the calculus of variations. If then this variation problem can be solved by one of the so-called direct methods, which do not make use of Euler equations, we have a possible method for the solution of the differential equation.

The direct method of Ritz[4] consists in starting with an independent and complete (but not necessarily orthogonal) system of given functions

$$\phi_1, \phi_2, \phi_3, \phi_4 \cdots$$

each exactly satisfying the boundary conditions in the domain being studied. In the integral

$$I = \int F\left(x, y, \cdots, \psi, \cdots \frac{\partial \psi}{\partial x}, \frac{\partial \psi}{\partial y}, \cdots\right) dxdy \ldots$$

to be minimized, one substitutes $\psi_n = \Sigma_1^n a_i \phi_i$, the a's being undetermined coefficients. The resulting integral, which we may call I_n, becomes a function of the a's and its minimization becomes an ordinary minimum problem. The a's are then determined by the relations $\partial I_n/\partial a_i = 0, i = 1, 2, \cdots n$. Then n is increased, and under ordinary conditions the sequence ψ_n approaches the solution ψ to the problem, while the corresponding sequence I_n approaches the minimum value of the given integral. One may, on physical grounds, be able to choose the initial terms of the functional sequence ϕ_i so they are near the real solution, thus securing rapid convergence.

Now the self-adjoint equation[5]

$$L[\phi] = (a\phi_x)_x + (b\phi_x)_y + (b\phi_y)_x + (c\phi_y)_y + d_x\phi + e_y\phi - f\phi = 0, \quad (1)$$

in which $\phi, a, b, \cdots, f$ are functions of x, y, is the most general homogeneous linear second order partial differential equation which is the Euler equation for a problem in the calculus of variations. In fact, if we consider solutions which are single valued, continuous and which vanish at the boundary Γ of a closed region G, then (1) is the Euler equation for a problem

$$\iint_G Q[\phi, \psi] dxdy = \min. \quad (2)$$

in which

$$Q[\phi\psi] = a\phi_x\psi_x + b\phi_x\psi_y + d\phi_y\psi_x + c\phi_y\psi_y + b\phi_x\psi + d\psi_x\phi + e\phi_y\psi + e\psi_y\phi + f\phi\psi.$$

Under the above conditions the extended Green formula becomes

$$\begin{aligned} \iint_G Q[\phi, \psi] &= -\iint_G \phi L[\psi] dxdy, \\ &= -\iint_G \psi L[\phi] dxdy; \end{aligned} \quad (3)$$

[4] W. Ritz, J. reine angew. Math. 135, 1–61. Collected works, pp. 192–250. Paris, 1911.

[5] The discussion of this paragraph is restricted for simplicity to the case of two independent variables. This restriction is not essential and the generalization is immediate.

for the usual integrations over Γ vanish.

The above statements assume that (2) actually possesses an extreme other than the trivial value zero. However, the homogeneous unconstrained problem does not possess such a minimum. Consider, however, the constrained problem

$$\int_G Q[\phi, \phi]d\tau = \min., \tag{4}$$

subject to the condition

$$\int_G \rho\phi^2 d\tau = 1, \tag{5}$$

where $d\tau$ is an element of coordinate space and ρ is a given function of the coordinates. By the ordinary methods of the calculus of variations the Euler equation in this case is

$$L[\phi] + e\rho\phi = 0. \tag{6}$$

When the integral (4) is minimized subject to (5), ϕ satisfies (6). Then by (3), (6) and (5) in turn,

$$\int Q[\phi, \phi]d\tau = -\int \phi L[\phi] = e\int \rho\phi^2 d\tau = e, \tag{7}$$

so that the extreme values of the integral are the characteristic values of (6).[6]

§2. *Extension to include perturbation effects.* Next we extend the above methods to enable us to investigate the effect upon the characteristic value e and the characteristic function ϕ of a small parameter f in L. Suppose

$$\begin{aligned} L[\phi] &= L_0[\phi] + fL_1[\phi] \\ \phi &= \phi_0 + f\phi_1 + f^2\phi_2 + \cdots \\ e &= e_0 + fe_1 + f^2e_2 + \cdots \end{aligned} \tag{8}$$

Then

$$O[\phi, \phi] = Q_0[\phi,\phi] + fQ_1[\phi,\phi].$$

Substitute (8) in (4) and (5) and collect powers of f. The zeroth order term furnishes[7]

$$-\int \phi_0 L_0[\phi_0]d\tau = \int Q_0[\phi_0,\phi_0]d\tau = \min. = e_0, \tag{9}$$

subject to

$$\int \rho\phi_0^2 d\tau = 1; \tag{10}$$

[6] For the material of this section, see Courant und Hilbert, Methoden der mathematischen Physik, Berlin, 1924.

[7] The minimum problem changes to a maximum problem because of the negative sign in Green's formula.

while from (6), ϕ_0 satisfies the Euler equation

$$L_0[\phi_0] + e_0\rho\phi_0 = 0. \tag{11}$$

The first order term furnishes

$$-\int (\phi_1 L_0[\phi_0] + \phi_0 L_0[\phi_1] + \phi_0 L_1[\phi_0])d\tau = \min = e_1, \tag{9'}$$

subject to

$$2\int \rho\phi_0\phi_1 d\tau = 0. \tag{10'}$$

Equation (9′) reduces by (3), (11), and (10′) to

$$-\int \phi_0 L_1[\phi_0]d\tau = \int Q_1[\phi_0\phi_0]d\tau = e_1; \tag{12'}$$

for since ϕ_0 is fixed by (9), there is no variation.

The quadratic terms in the parameter f furnish

$$-\int (\phi_2 L_0[\phi_0] + \phi_1 L_0[\phi_1] + \phi_0 L_0[\phi_2] + \phi_1 L_1[\phi_0] + \phi_0 L_1[\phi_1])d\tau \tag{9''}$$

$$= \min = e_2,$$

subject to

$$\int \rho(\phi_1^2 + 2\phi_0\phi_2)d\tau = 0; \tag{10''}$$

and 9′′ becomes by (3), (11), and (10′′),

$$\begin{aligned} &-\int (e_0\rho\phi_1^2 + \phi_1 L_0[\phi_1] + 2\phi_1 L_1[\phi_0])d\tau, \\ &= \int (Q_0[\phi_1\phi_1] + 2Q_1[\phi_1\phi_0] - e_0\rho\phi_1^2)d\tau = \min = e_2. \end{aligned} \tag{12''}$$

Hence we see that by solving (9) subject to the constraint (10) and then (12′′) subject to (10′), one obtains e_0, ϕ_0, e_2 and ϕ_1, while e_1 is directly obtained from the integral (12′).

B. Preliminaries to Calculation

Units and the Schrödinger equation. If one uses a unit of length $a_H/2Z = h^2/8\pi^2 mZe'^2$, a unit of energy Z^2Rh and a unit of charge e', the field θ will, following its customary definition, be measured in terms of the unit $16\pi^4 m^2 Z^3 e'^5/h^4$. Then the wave equation for a helium atom in an electric field is

$$(\Delta_1 + \Delta_2)\psi + \left(\frac{\lambda}{4} + \frac{1}{r_1} + \frac{1}{r_2} - \frac{1}{2r_{12}} + \frac{\theta}{4}(z_1 + z_2)\right)\psi = 0$$

Here λ is the energy parameter; r_1, r_2, r_{12}, z_1, z_2 are distances of the electrons from the nucleus, distance between the electrons, and z coordinates of the electron measured parallel to the field. For further simplicity we make the substitutions $\lambda/4=e$ and $\theta/4=f$ which reduce the Schrödinger equation to

$$(\Delta_1+\Delta_2)\psi+(e+1/r_1+1/r_2-1/2r_{12}+f(z_1+z_2)\psi=0$$

Let E be the energy of an atom when placed in an electric field F. Then the expansion

$$E=E_0+E_1F+E_2F^2+E_3F^3+\cdots$$

is useful in the discussion of the dielectric properties of the atom. This equation may be thought of as being written in ordinary units, while in the special units here employed,

$$e=e_0+e_1f+e_2f^2+\cdots$$

If the units of this equation are to be consistent we must have e_2 measured in a unit $h^6/512\pi^6m^3z^4e'^6$ times as large as the unit of E_2.

The dielectric constant. If ϵ is the dielectric constant and N is the number of molecules per unit volume at the temperature and pressure at which ϵ is measured, it is true that

$$\epsilon-1=-8\pi NE_2,$$

or

$$\epsilon-1=-\frac{e_2h^6N}{64\pi^5z^4e'^6m^3}$$

in the units here employed. An experimental value[9] for $\epsilon-1$ is 0.000070 and this requires e_2 to be -88 (at 0° C and 760 mm of Hg).

Coordinates to be used. Hylleraas (II) has shown the great advantage of using the distances which occur in the potential terms as coordinates. For unperturbed helium he used as coordinates the distances r_1, r_2, r_{12}, the three angles which might be added to these distances to make a complete set of coordinates being ignorable. In our case we may use as coordinates r_1, r_2, r_{12}, z_1, z_2, the one other azimuthal coordinate being ignorable because of the symmetry of the problem.

Six dimensional Laplace operator in coordinates r_1, r_2, r_{12}, z_1, z_2. Suppose $\psi=\psi(r_1, r_2, r_{12}, z_1, z_2)$, in which

$$r_1^2=x_1^2+y_1^2+z_1^2,\quad r_2^2=x_2^2+y_2^2+z_2^2,$$
$$r_{12}^2=(x_1-x_2)^2+(y_1-y_2)^2+(z_1-z_2)^2,$$
$$z_1=z_1,\quad z_2=z_2.$$

Then $(\Delta_1+\Delta_2)\psi=$

$$\frac{\partial^2\psi}{\partial r_1^2}+\frac{2}{r_1}\frac{\partial\psi}{\partial r_1}+\frac{\partial^2\psi}{\partial r_2^2}+\frac{2}{r_2}\frac{\partial\psi}{\partial r_2}+2\frac{\partial^2\psi}{\partial r_{12}}+\frac{4}{r_{12}}\frac{\partial\psi}{\partial r_{12}}$$

[8] J. H. Van Vleck, Proc. Nat. Acad. **12**, 663 (1926).

[9] Herzfeld and Wolf, Ann. d. Physik **76**, 71 and 567, find $\epsilon-1=0.0000693$ from extrapolation of optical data. Direct measurement usually gives results somewhat higher. C. P. Hochhein, Verh. D. Phys. Ges. **10**, 446 (1908): $\epsilon-1=0.000074$.

$$+\frac{r_1^2-r_2^2+r_{12}^2}{r_1r_{12}}\frac{\partial^2\psi}{\partial r_1\partial r_{12}}+\frac{r_2^2-r_1^2+r_{12}^2}{r_2r_{12}}\frac{\partial^2\psi}{\partial r_2\partial r_{12}}+\frac{2z_1}{r_1}\frac{\partial^2\psi}{\partial r_1\partial z_1}$$
$$+\frac{2(z_1-z_2)}{r_{12}}\frac{\partial^2\psi}{\partial r_{12}\partial z_1}+\frac{2z_2}{r_2}\frac{\partial^2\psi}{\partial r_2\partial z_2}+\frac{2(z_2-z_1)}{r_{12}}\frac{\partial^2\psi}{\partial r_{12}\partial z_2}$$
$$+\frac{\partial^2\psi}{\partial z_1^2}+\frac{\partial^2\psi}{\partial z_2^2}.$$

Jacobian. The differential expression just written is not self-adjoint. We can make it so by multiplication by the Jacobian of the transformation from rectangular coordinates to the coordinates $r_1, r_2, r_{12}, z_1, z_2, \chi$ in which χ is some angle that has been added to make the coordinate system complete. If one computes this Jacobian and then eliminates the ignorable coordinate χ by integration over it, the result is, aside from trivial constant factors,

$$\rho=\frac{r_{12}r_1r_2}{q^{1/2}},$$

in which

$$-q=4r_1^2z_2^2+(4r_{12}^2-4r_1^2-4r_2^2)z_1z_2+4r_2^2z_1^2$$
$$+r_{12}^4-2r_{12}^2r_1^2+r_1^4-2r_{12}^2r_2^2-2r_1^2r_2^2+r_2^4.$$

To simplify connection with Hylleraas II on which we shall depend for the unperturbed characteristic functions, we make the further change of the variable

$$s=r_1+r_2,$$
$$t=-r_1+r_2,$$
$$u=r_{12}.$$

The Jacobian of this last transformation is a constant so that the Jacobian weight factor of the combined transformation is essentially

$$\sigma=\frac{(s^2-t^2)u}{q^{1/2}},$$

in which

$$-q=(s-t)^2z_2^2+(4u^2-2s^2-2t^2)z_1z_2+(s+t)^2z_1^2$$
$$+u^4-u^2(s^2+t^2)+s^2t^2.$$

We omit the differential expressions for $(\Delta_1+\Delta_2)\psi$ in the coordinates s, t, u, z_1, z_2, but give rather the formulation of the problem in the calculus of variations

$$Q_0[\psi,\psi]=P[\psi,\psi]-\sigma\left(\frac{2}{s-t}+\frac{2}{s+t}-\frac{1}{2u}\right)\psi^2,$$
$$Q_1[\phi,\psi]=-\sigma(z_1+z_2)\phi\psi,$$

in which

$$P[\psi,\psi] = 2\sigma\psi_s^2 + 2\sigma\psi_t^2 + 2\sigma\psi_u^2 + 4\sigma\left\{\frac{s(u^2-t^2)}{u(s^2-t^2)}\psi_s + \frac{t(s^2-u^2)}{u(s^2-t^2)}\psi_t\right\}\psi_u + 4\sigma\frac{z_1}{s-t}(\psi_s-\psi_t)\psi_{z_1} + 2\sigma\frac{z_1-z_2}{u}\psi_u\psi_{z_1} + 4\sigma\frac{z_2}{s+t}(\psi_s+\psi_t)\psi_{z_2} + 2\sigma\frac{z_2-z_1}{u}\psi_u\psi_{z_1} + \sigma\psi_{z_1}^2 + \sigma\psi_{z_2}^2.$$

Limits of integration. If $d\tau$ is an element of coordinate space, i.e., if $d\tau = dsdtdudz_1dz_2$, then

$$\int d\tau = \int_0^\infty ds \int_0^s du \int_0^u dt \int_{-(s-t)/2}^{(s-t)/2} dz_1 \int_A^B dz_2,$$

where A and B are roots of $q=0$. These limits only cover part of space in which $r_2 > r_1$. Now in the normal state the wave function is symmetrical in the two electrons and so the limits just written produce one-half of the total integral. This factor of one-half is, however, easily compensated for by using the same limits in the normalization.

The integrals involved. The fact that the limits A and B are zeros of q or infinities of σ is of importance in performing the integrations over the z coordinates. A and B are clearly branch points of the integrand. In general, they will be the only singular points in the finite z_2 plane. So one easily finds by integration in the complex plane

$$\int_{-(s-t)/2}^{(s-t)/2}\int_A^B \sigma z_1^2 dz_1dz_2 = \frac{\pi}{12}u(s-t)^3(s+t)$$

$$\int_{-(s-t)/2}^{(s-t)/2} \sigma z_1z_2 dz_1dz_2 = \frac{\pi}{12}u(s^2-t^2)(s^2+t^2-2u^2)$$

$$\int_{-(s-t)/2}^{(s-t)/2}\int_A^B \sigma z_2^2 dz_1dz_2 = \frac{\pi}{12}u(s+t)^3(s-t)$$

$$\int_{-(s-t)/2}^{(s+t)/2}\int_A^B \sigma dz_1dz_2 = \pi u(s^2-t^2).$$

The last result should be the weight factor when there are no z coordinates. Hylleraas II gives the weight factor

$$\frac{u(s^2-t^2)}{8}$$

for this case, showing that his normalization differs from ours by a factor 8π.

The other integrals that we shall need to evaluate are all of the form

$$\int_0^\infty ds \int_0^s du \int_0^u dt e^{-as} s^p t^q u^r = \frac{1}{a^{p+q+r+3}} \frac{(p+q+r+2)!}{(q+1)(r+q+2)}.$$

C. Actual Calculation of Dielectric Constant

We are now prepared to carry out the actual minimization of

$$I = \int \{Q_0[\psi_1,,\psi_1] + 2Q_1[\psi_1,\psi_0] - e_0\sigma\psi_1^2\} d\tau = \min = e_2,. \tag{1}$$

when ψ_0 is given, subject to

$$\int \sigma\psi_0\psi_1 d\tau = 0. \tag{2}$$

The function ψ_1 must, of course, satisfy all the conditions imposed on any solution of the Schrödinger equation. But, in addition, it must be symmetrical in the coordinates of the two electrons for it forms part of the wave function for a par-helium state. This requires, since $s=r_1+r_2$ and $t=-r_1+r_2$ that $\psi_1(s, t, u, z_1, z_2)=\psi_1(s,-t, u, z_2, z_1)$. The actual choosing of the functional form will depend on physical knowledge and experience with previous trials. For the form of Q_0 and Q_1, see page 1237.

First approximation. It is easily shown when this same method is applied to hydrogen that a very simple wave function of the form

$$\psi_1 = \alpha_1 e^{-r/2} z$$

gives results for $-e_2$ that are only 11% low. By ideas which will be given later we easily generalize this to apply to helium thus obtaining[10]

$$\psi_1 = \alpha_1 e^{-c_1 s/2}(z_1 + z_2). \tag{3}$$

If one substitutes this in the expression (1) to be minimized and uses as ψ_0

$$\begin{aligned} \psi_0 &= e^{-c_1 s/2}(1 + c_2 u + c_3 t^2), \\ c_1 &= 0.908, \\ c_2 &= 0.0726, \\ c_3 &= 0.0083; \end{aligned} \tag{4}$$

(one of the results of Hylleraas II which gives the normal energy level to within 1/30%) one obtains on integration

$$I = \frac{\pi}{6c_1^6}\{391.1\alpha_1^2 - 5787\alpha_1\} = \min.$$

Now α_1 for a minimum is 7.40, so that

[10] Professor Van Vleck has suggested that the most general form of ψ_1 is $\psi_1=z_1 f(s,t,u) + z_2 f(s,-t,u)$. It is easily seen that this form satisfies the non-homogeneous wave equation for the perturbing function.

$$I = \frac{\pi}{c_1^6}(-3{,}568).$$

But ψ_0 is not normalized. We divide I by the normalization factor[11]

$$\frac{\pi}{c_1^6}65.587,$$

and then find

$$-e_2 = 54.4$$

a value 37 percent low as compared with the accepted value $-e_2=88$ (page 1236.

Second approximation. We know that for normal hydrogen in our coordinates

$$\psi_0 = \tfrac{1}{2}e^{-r/2},$$
$$\psi_1 = \tfrac{1}{4}z(r-4)e^{-r/2},$$

and consequently to a second approximation

$$\psi = \psi_0 + f\psi_1 = e^{-r/2}\left\{\frac{1}{2} + f\frac{z}{4}(r-4)\right\}.$$

If one neglects interactions between the electrons in a helium atom, the Schrödinger equation becomes separable and the true wave function for the perturbed problems becomes

$$\psi = e^{-r_1/2-r_2/2}\left\{\frac{1}{2} + f\frac{z_1}{4}(r_1+4)\right\}\left\{\frac{1}{2} + f\frac{z_2}{4}(r_2+4)\right\}.$$

Thus

$$\psi_1 = e^{-s/2}\left\{z_1\left(\frac{r_1}{8}+\frac{1}{2}\right) + z_2\left(\frac{r_2}{8}+\frac{1}{2}\right)\right\},$$

for ψ_1 is the coefficient of f in the power development of ψ. It is natural to generalize this into[11]

$$\psi_1 = e^{-c_1 s/2}\left\{z_1(\alpha_1 + \alpha_2 s - \alpha_3 t) + z_2(\alpha_1 + \alpha_2 s + \alpha_3 t)\right\} \qquad (5)$$

in which α_1, α_2, and α_3 are unknown constants. We choose for c_1 the value that Hylleraas used in formula (4) (i.e., $c_1=0.908$). One notices that this function remains unchanged by interchanging the electrons, and that it is orthogonal to the ψ_0 given in (4), i.e.,

$$\int \sigma\psi_0\psi_1 d\tau = 0.$$

We substitute this expression (5) in the integral (1) to be minimized using (4) as ψ_0, and obtain after carrying out the integrations

[11] The normalization factors are most easily obtained by multiplication of Hylleraas' factors by 8π due to difference in normalization and division by c_1^6 to take care of his change in scale.

$$I = \frac{\pi}{c_1^6}(65.2\alpha_1^2 + 485.8 \times 2\alpha_1\alpha_2 + 4770 \times \alpha_2^2 + 38 \times 2\alpha_1\alpha_3 + 303 \times 2\alpha_1\alpha_3 + 979 \times \alpha_3^2 - 479 \times 2\alpha_1 - 4447 \times 2\alpha_2 - 1287 \times 2\alpha_3) = \text{min.}$$

To find the minimum, we solve the linear equations

$$\frac{\partial I}{\partial \alpha_1} = 0$$

$$\frac{\partial I}{\partial \alpha_2} = 0$$

$$\frac{\partial I}{\partial \alpha_3} = 0$$

and obtain

$$\alpha_1 = 1.87, \quad \alpha_2 = 0.673, \quad \alpha_3 = 1.03.$$

On substitution of these values in the quadratic expression we find

$$I = 5091\frac{\pi}{c_1^6}.$$

So $-e_2 = \dfrac{5091(\pi/c_1^6)}{65.59(\pi/c_1^6)} = 78$. This is 11 percent below the experimental value of $-e_2 = 88$.

Third approximation. We now extend the function ψ_1 by the addition of a term in u; α_1, α_2, α_3, α_4 being new unknown constants

$$\psi_1 = e^{-c_1 s/2}\{(\alpha_1 + \alpha_2 s - \alpha_3 t + \alpha_4 u)z_1 + (\alpha_1 + \alpha_2 s + \alpha_3 t + \alpha_4 u)z_2\}.$$

It now seems worth while to employ for ψ_0 the final and best result Hylleraas obtained in his second paper. This is

$$\psi_0 = e^{-c_1 s/2}(1 + c_2 u + c_3 t^2 + c_4 s + c_5 s^2 + c_6 u^2),$$

$$\begin{aligned} c_1 &= 0.9100, \\ c_2 &= 0.0885, \\ c_3 &= 0.0080, \\ c_4 &= -0.0252, \\ c_5 &= 0.0021, \\ c_6 &= -0.0020. \end{aligned} \tag{7}$$

This wave function is sufficiently accurate to give the normal energy level of helium to within 0.01 percent of its experimental value.

On substitution of these expressions and integrating, the following quadratic expression is obtained

$$I = \frac{\pi}{c_1^6}[65.14\alpha_1^2 + 483.39 \times 2\alpha_1\alpha_2 + 4716.1\alpha_2^2 + 37.59 \times 2\alpha_1\alpha_3 + 292.7 \times 2\alpha_2\alpha_3 + 974.3\alpha_3^2 + 269.58 \times 2\alpha_1\alpha_4 + 2598.3 \times 2\alpha_2\alpha_4 + 499.3 \times 2\alpha_3\alpha_4 + 1922.5\alpha_4^2 - 465.46 \times 2\alpha_1 - 4336.4 \times 2\alpha_2 - 1240.6\alpha_3 - 2960.1 \times 2\alpha_4] = \min.$$

Minimizing as before yields

$$\alpha_1 = 0.825, \quad \alpha_2 = 0.456, \quad \alpha_3 = 0.815, \quad \alpha_4 = 0.596$$

$$I = \frac{\pi}{c_1^6}(-5124).$$

The normalization factor is now

$$\frac{\pi}{c_1^6}61.496.$$

This yields $-e_2 = 83.4$ which is less than 5 percent below the experimental value of $-e_2 = 88$.

It remains the author's pleasant duty to express his appreciation to Professor J. H. Van Vleck, to Professor G. Wentzel, and to Professor Warren Weaver for much helpful advice given throughout the preparation of this paper.

Note added in proof. The quantity e_2 is given by equation C(1) as the minimum value of an integral with respect to the variations of an unknown (and thus determined) function ψ_1, the function ψ_0 being assumed known. Prof. J. H. Van Vleck has kindly suggested that one should emphasize the following consequence of this: That first order errors in ψ_1 will only produce second order errors in e_2 but that on the contrary first order errors in ψ_0 will likely produce first order errors in e_2. This shows mathematically the importance of an accurate knowledge of the umperturbed state in the calculation of the dielectric constant.

THE POTENTIAL AND POTENTIAL ENERGY OF SPACE LATTICES

By C. N. Wall
North Central College, Naperville, Illinois

(Received July 24, 1930)

Abstract

The electrostatic potential of a general space lattice is developed. The space lattice is characterized by a base cell containing a finite set of positive point charges arbitrary in strength and position, and a negative space charge of arbitrary density, subject to the condition that the total charge in the cell is zero.

Next the expression for the lattice energy is obtained in the form of a triply infinite series. It is shown that the coefficients in this series representing the distribution of the negative space charge can be identified with the structure factors of the lattice or crystal provided we replace the negative space charge by a corresponding electron distribution.

An application of this theory is made to the three halides NaCl, NaF, and LiF. The lattice energy of each crystal is calculated for different grating spaces. In all three cases it is shown that the lattice energy has a minimum in the neighborhood of the accepted grating space for the crystal under consideration. The agreement is better for NaCl than for NaF or LiF.

The Electrostatic Potential

THE electrostatic potential of an infinite space lattice with an arbitrary distribution of point charges in the base cell can be determined by a method due to Ewald.[1] This method consists in considering a continuous periodic distribution of charges the density of which can be represented by a triple Fourier series without a constant term.[2] The potential of such a system can be represented by a similar Fourier series whose coefficients can be determined by the use of Poisson's equation. Finally one arrives at the required result by taking the limiting case in which the continuous space charge shrinks into a discrete set of point charges. If, in this process, we allow the positive space charge to shrink into point charges but keep the negative charge as a continuous distribution of electricity, we obtain a space lattice with cells consisting of a set of positive point charges (nuclei) surrounded by an atmosphere of negative electricity which may be identified with the electron distribution under certain conditions. It is the potential and potential energy of such a space lattice which we wish to determine.

Let us consider the base cell of our lattice to be defined by the three vectors a^1, a^2, a^3. Our lattice can be built up by a simple translation of this base cell in three space directions defined by a^1, a^2, a^3. The lattice shall be referred to the rectangular axes x_1, x_2, x_3 whose origin shall be at a vertex

[1] P. P. Ewald, Ann. d. Physik **64**, 253 (1921).
M. Born, Problems of Atomic Dynamics, 158–162 (1926).

[2] The lack of a constant term in the series means that the total charge in any cell of the lattice is zero.

of the base cell. The rectangular components of the three base vectors shall be designated by a_i^m; $m, i = 1, 2, 3$. a_i^m is the x_ith component of the a^mth vector. The volume of the base cell shall be $\Delta = |a_i^m|$. We introduce the three vectors of the reciprocal lattice b^1, b^2, b^3 with their components b_j^n defined by the equations

$$\sum_{i=1}^{3} a_i^m b_i^n = \left\{ \begin{matrix} 0 \text{ if } n \neq m \\ 1 \text{ if } n = m \end{matrix} \right\}; \; m, n = 1, 2, 3.$$

Let there be p positive point charges in the base cell, the kth charge having a strength e_k and coordinates x_i^k. The negative space charge shall have a density distribution of $\rho(x_1, x_2, x_3)$ in the base cell. We impose the condition that

$$\sum_{k=1}^{p} e_k + \iiint \rho dx_1 dx_2 dx_3 = 0$$

where the integral is taken throughout the entire base cell. This condition is equivalent to the statement that the total charge in the base cell, and therefore in any cell, is zero.

With this brief description of the lattice under consideration, we write down at once the potential of this lattice as given by Ewald's method.

$$V = \frac{1}{\pi\Delta} \sum_{l_1, l_2, l_3 = -\infty}^{\infty}{}' \left\{ \frac{A_{l_1 l_2 l_3} + B_{l_1 l_2 l_3}}{\sum_{j=1}^{3} \left(\sum_{n=1}^{3} l_n b_j^n \right)^2} \right\} \exp\left(i2\pi \sum_{j,n=1}^{3} l_n b_j^n x_j \right) \tag{1}$$

where the coefficients $A_{l_1 l_2 l_3}$ and $B_{l_1 l_2 l_3}$ are given by the equations

$$\left\{ \begin{aligned} A_{l_1 l_2 l_3} &= \sum_{k=1}^{p} e_k \exp\left(- i2\pi \sum_{j,n=1}^{3} l_n b_j^n x_j^k \right) \\ B_{l_1 l_2 l_3} &= \iiint \rho(\xi_1, \xi_2, \xi_3) \exp\left(- i2\pi \sum_{j,n=1}^{3} l_n b_j^n \xi_j \right) d\xi_1 d\xi_2 d\xi_3 \end{aligned} \right\} \tag{2}$$

the integration to be taken throughout the base cell. The coefficients $A_{l_1 l_2 l_3}$ are functions of the positions and the strengths of the positive point charges, and the coefficients $B_{l_1 l_2 l_3}$ are functions of the distribution of the negative space charge. The total potential V is obviously the sum of the separate potentials of the positive and the negative charges, but the above series written with all of the $B_{l_1 l_2 l_3}$ coefficients omitted does not represent the potential of the positive charges alone. Rather it gives the potential of the positive charges imbedded in a negative space charge of uniform density such that the total charge in any cell is zero. It is seen by Eqs. (2) that for $\rho = \text{constant}$, $B_{l_1 l_2 l_3} = 0$ except for $l_1 = l_2 = l_3 = 0$, and this term is omitted in the series in Eq. (1) as indicated by the prime on the summation sign.

The Potential Energy

Having calculated the potential of the space lattice under consideration we turn to the problem of determining the potential energy of the lattice,

or rather the potential energy of a single cell of the lattice. For this calculation we make use of Green's theorem which may be written in the form

$$\iiint V\nabla^2 V\,dx_1dx_2dx_3 + \iiint [(\partial V/\partial x_1)^2 + (\partial V/\partial x_2)^2 + (\partial V/\partial x_3)^2]dx_1dx_2dx_3 + \sum \iint V(\partial V/\partial n)dS = 0$$

where the symbols have their customary significance.

The potential energy of the base cell will be given by the expression

$$(1/8\pi)\iiint [(\partial V/\partial x_1)^2 + (\partial V/\partial x_2)^2 + (\partial V/\partial x_3)^2]dx_1dx_2dx_3$$

where the integration is taken throughout the base cell. However this expression will give an infinite energy because of the presence of positive point charges in the cell of infinite self-energy. We can overcome this difficulty by deleting each of the positive charges with a small sphere of radius δ and subtracting off the self-energy of these positive charges. The resultant potential energy of the cell which we shall designate by Φ will then be given by the equation

$$\Phi = \lim_{\delta\to 0}\left\{(1/8\pi)\iiint [(\partial V/\partial x_1)^2 + (\partial V/\partial x_2)^2 + (\partial V/\partial x_3)^2]dx_1dx_2dx_3 - \sum_{k=1}^{p}(e_k)^2/2\delta\right\} \quad (3)$$

where the integration now extends through-out the base cell excluding the p small spheres of radii δ.

By the use of Green's theorem Φ can be written in the form

$$\Phi = \lim_{\delta\to 0}\left\{-(1/8\pi)\iiint V\nabla^2 V\,dx_1dx_2dx_3 - (1/8\pi)\sum\iint V(\partial V/\partial n)dS - \sum_{k=1}^{p}(e_k)^2/2\delta\right\}. \quad (4)$$

Now $\nabla^2 V = -4\pi\rho$ and $\iint V(\partial V/\partial n)dS$ vanishes over the external surface of the cell because of the periodicity of V and $\partial V/\partial n$.[3] Thus the surface integral of $V\partial V/\partial n$ reduces to that over the small spheres alone. Eq. (4) may now be written

[3] If some of the point charges lie on the bounding surface of the cell, they may be deleted by small hemispherical indentations precisely as is done in the theory of elliptic functions.

$$\Phi = \lim_{\delta \to 0} \Big\{ (1/2) \iiint \rho V dx_1 dx_2 dx_3 - (1/8\pi) \sum_{k=1}^{p} \iint V(\partial V/\partial n) dS - \sum_{k=1}^{p} (e_k)^2/2\delta \Big\}. \tag{5}$$

But $\lim_{\delta \to 0}(1/2)\iiint \rho V dx_1 dx_2 dx_3$ is equal to $(1/2)\iiint \rho V dx_1 dx_2 dx_3$, the integral being taken through-out the entire base cell, since the integral of ρV throughout any of the small spheres goes to zero as δ goes to zero. Also it is not difficult to show that

$$\lim_{\delta \to 0} \Big\{ - (1/8\pi) \sum_{k=1}^{p} \iint V(\partial V/\partial n) dS - \sum_{k=1}^{p} (e_k)^2/2\delta \Big\} = (\tfrac{1}{2}) \sum_{k=1}^{p} e_k V_k'$$

where V_k' is the potential at the point charge e_k with that charge removed. Making these substitutions in Eq. (5) we obtain

$$\Phi = (\tfrac{1}{2}) \iiint \rho V dx_1 dx_2 dx_3 + (\tfrac{1}{2}) \sum_{k=1}^{p} e_k V_k'. \tag{6}$$

Substituting in Eq. (6) the value of V as given in Eq. (1) we get

$$\Phi = (1/2\pi\Delta) \sum_{l_1, l_2, l_3 = -\infty}^{\infty}{}' \frac{(A_{l_1 l_2 l_3} + B_{l_1 l_2 l_3}) B_{-l_1 -l_2 -l_3}}{\sum_{j=1}^{3} \left(\sum_{n=1}^{3} l_n b_j{}^n \right)^2} + (\tfrac{1}{2}) \sum_{k=1}^{p} e_k V_k'. \tag{7}$$

Let us consider the quantity $(1/2)\sum_{k=1}^{p} e_k V_k'$ appearing in Eq. (7). V_k' is the potential at the k th positive charge with that charge removed. It is made up of two parts, that due to all the other positive charges, and that due to the negative space charge. Calling the former $V_k'{}_+$ and the latter $V_k'{}_-$ we see by the use of Eq. (1) that

$$(\tfrac{1}{2}) \sum_{k=1}^{p} e_k V_k' = (\tfrac{1}{2}) \sum_{k=1}^{p} e_k (V_k'{}_+ + V_k'{}_-)$$

$$= (\tfrac{1}{2}) \sum_{k=1}^{p} e_k V_k'{}_+ + (1/2\pi\Delta) \sum_{l_1, l_2, l_3 = -\infty}^{\infty}{}' \frac{B_{l_1 l_2 l_3} A_{-l_1 -l_2 -l_3}}{\sum_{j=1}^{3} \left(\sum_{n=1}^{3} l_n b_j{}^n \right)^2}. \tag{8}$$

Substituting in Eq. (7) the results expressed in Eq. (8) and collecting terms we obtain Φ in its final form

$$\Phi = (1/2\pi\Delta) \sum_{l_1, l_2, l_3 = -\infty}^{\infty}{}' \frac{B_{l_1 l_2 l_3} B_{-l_1 -l_2 -l_3} + 2A_{l_1 l_2 l_3} B_{-l_1 -l_2 -l_3}}{\sum_{j=1}^{3} \left(\sum_{n=1}^{3} l_n b_j{}^n \right)^2} + (\tfrac{1}{2}) \sum_{k=1}^{p} e_k V'_{k+}. \tag{9}$$

Eq. (9) gives in compact form the potential energy of the base cell of the lattice under consideration. Given the quantities $a_j{}^n$, e_k, $x_i{}^k$, and the func-

tion $\rho(x_1, x_2, x_3)$, the value of Φ may be calculated by Eq. (9). A method devised by Ewald enables us to calculate the quantity $(1/2)\sum_{k=1}^{p} e_k V_k'_+$ without too much difficulty.[4]

An examination of Eq. (9) shows that the potential energy of the cell consists of three parts. The first part is the energy of the negative space charge and involves the coefficients $B_{l_1 l_2 l_3}$ $B_{-l_1-l_2-l_3}$. The second part is the mutual energy of the negative space charge and the positive point charges. It involves the coefficients $A_{l_1 l_2 l_3}$ $B_{-l_1-l_2-l_3}$. The third part as given by $(1/2)\sum_{k=1}^{p} e_k V_k'_+$ is the energy of the positive point charges. The first and third parts are essentially positive while the second part is negative.

The expression for Φ as given in Eq. (9) has some advantages over the expressions usually given for the potential energy of a lattice cell. It is usually assumed that the ions making up an ionic crystal can be treated as point charges. No attempt is made to take into account the actual distribution of electrons around the nuclei. As a result it is necessary, for the sake of achieving equilibrium, to introduce an additive term in the energy expression representing the effect of certain repulsive forces existing between the sions. In the expression here developed we have taken into consideration the distribution of the electrons in the crystal in so far as it is possible to replace a negative space charge by a corresponding electron atmosphere. This obviates the necessity of introducing extra terms representing repulsion since the sheath of electrons around each nuclei will automtically bring into play repulsive forces of considerable magnitude when the ions are close together.

Furthermore the coefficients $B_{l_1 l_2 l_3}$ which play an important part in this theory can, with only a few minor assumptions, be identified with the structure factors of the crystal in question.[5] The structure factors for some crystals have been determined experimentally. For these crystals Φ may be calculated not only for the accepted crystal parameters but also for arbitrary parameters. This enables us to determine parameter values which give minimum values of Φ and therefore states of stable equilibrium. In the following section we give the results of some determinations of Φ for the halides NaCl, NaF, and LiF.

Applications

The halides NaCl, NaF, and LiF are simple cubic crystals. The expression for Φ as given in Eq. (9) becomes considerably simpler for this type of a crystal. We carry through the computation of Φ for NaCl as illustrating the general method. In order to use the experimentally determined values of the structure factors it is necessary to take the origin at the heaviest ion in the crystal.[6] This is at the Cl ion for NaCl.

There are two kinds of ions in the NaCl crystal so that

[4] e. g. M. Born, 159–162.

[5] A. H. Compton, X-Rays and Electrons, Chap. 5 (1926).

[6] R. J. Havighurst, Phys. Rev. **29**, 4 (1927).

$$e_1 = e_3 = e_5 = e_7 = 17e, \quad \text{and} \quad e_2 = e_4 = e_6 = e_8 = 11e$$

where e is numerically equal to the charge on the electron. Since NaCl is a cubic crystal we have

$$a_j^n = \begin{cases} 0 & \text{if } n \neq j \\ a & \text{if } n = j \end{cases}, \quad \text{and} \quad b_j^n = \begin{cases} 0 & \text{if } n \neq j \\ b & \text{if } n = j \end{cases}.$$

Evidently $a = 1/b$ and $\Delta = a^3$. There are eight ions in each cell so that k runs from 1 to 8. The coordinates of the nuclei of these ions are

$$x_1^1 = x_2^1 = x_3^1 = x_2^2 = x_3^2 = x_2^3 = x_1^4 = x_3^4 = x_1^5 = x_1^6 = x_2^6 = x_3^7 = 0,$$

and

$$x_1^2 = x_1^3 = x_3^3 = x_2^4 = x_2^5 = x_3^5 = x_3^6 = x_1^7 = x_2^7 = x_1^8 = x_2^8 = x_3^8 = a/2.$$

We consider $\rho(\xi_1, \xi_2, \xi_3)$ as an even function of its arguments so that

$$B_{l_1 l_2 l_3} = \iiint \rho \exp\left[-i2\pi b(l_1\xi_1 + l_2\xi_2 + l_3\xi_3)\right] d\xi_1 d\xi_2 d\xi_3$$

$$= \iiint \rho \cos\left[2\pi b(l_1\xi_1 + l_2\xi_2 + l_3\xi_3)\right] d\xi_1 d\xi_2 d\xi_3 = B_{\pm l_1 \pm l_2 \pm l_3}.$$

So far we have considered ρ to be the negative space charge density in the crystal. We assume that it may be set equal to $-e\sigma$ where σ is the electron density in the crystal. But

$$\iiint \sigma \cos\left[2\pi b(l_1\xi_1 + l_2\xi_2 + l_3\xi_3)\right] d\xi_1 d\xi_2 d\xi_3 = 4F_{l_1 l_2 l_3}$$

where $F_{l_1 l_2 l_3}$ is the l_1, l_2, l_3 structure factor for the NaCl molecule in the crystal.[7] The factor 4 enters because there are four molecules in each cell of the crystal. Thus we have

$$B_{\pm l_1 \pm l_2 \pm l_3} = -4eF_{l_1 l_2 l_3}.$$

The expression for Φ with the above modifications now takes the form

$$\Phi = (\tfrac{1}{2}) \sum_{k=1}^{8} e_k V_k' + (1/2\pi a) \sum_{l_1, l_2, l_3 = 0}^{\infty}{}' \delta_{l_1 l_2 l_3} \left(\frac{-8e\alpha_{l_1 l_2 l_3} F_{l_1 l_2 l_3} + 16e^2 (F_{l_1 l_2 l_3})^2}{l_1^2 + l_2^2 + l_3^2} \right) \quad (10)$$

where

$$\delta_{l_1 l_2 l_3} = \begin{cases} 8 \text{ if none of the subscripts are zero,} \\ 4 \text{ if only one of the subscripts is zero,} \\ 2 \text{ if only two of the subscripts are zero} \end{cases},$$

and where

$$\alpha_{l_1 l_2 l_3} = \begin{cases} 112e & \text{for the subscripts all even} \\ 24e & \text{for the subscripts all odd} \\ 0 & \text{otherwise} \end{cases}.$$

[7] A. H. Compton, p. 160.

We are now in a position to calculate Φ for NaCl by the use of Eq. (10). The calculation of the term $(1/2)\sum_{k=1}^{8} e_k V_k^1{}_+$ is a rather long and laborious process. The standard method due to Ewald[8] is used for this calculation and need not be repeated here. The result for NaCl is $355e^2/a$. This amounts to less than three percent of the total value of Φ and thus contributes little to the energy of the cell.

The values of the structure factors for NaCl, and also for NaF and LiF, have been taken from a table of structure factors given by R. J. Havighurst.[6] These values as given by Havighurst include the Debye temperature factor. Using forty of these values and substituting in Eq. 10 we obtain for Φ of NaCl the value

$$\Phi\ (NaCl) \equiv -\ 13590\ e^2/a_0$$

where $a_0 = 5.628 \times 10^{-8}$ cm. This value of Φ cannot be compared with Born's value of the potential energy of a NaCl cell since this value includes the energy of the ions and takes into account the electron distribution in the crystal.

It is possible to compute Φ (NaCl) for various arbitrary values of the grating space since the structure factors for various values of a can be determined. The same can be done for NaF and LiF. The results are shown in tabulated form. a_0 represents the accepted grating space for the crystals under consideration and has the values $a_0(\text{NaCl}) = 5.628 \times 10^{-8}$ cm., $a_0(\text{NaF}) = 4.620 \times 10^{-8}$ cm, and $a_0(\text{LiF}) = 4.014 \times 10^{-8}$ cm. Values of $w = (a_0/e^2)\Phi$ are shown for the grating spaces $0.75a_0$, a_0, $1.25a_0$, $1.50a_0$, and $2.00a_0$.

TABLE I

	$0.75a_0$	a_0	$1.25a_0$	$1.50a_0$	$2.00a_0$
w(NaCl)	−11860	−13590	−12850	—	—
w(NaF)	− 4350	− 5360	− 5470	−4850	
w(LiF)[9]	− 1710	− 1950	—	−1980	−1800

An examination of Table I yields the interesting fact that a minimum value of w, and therefore of Φ, occurs for each of the three crystals in the neighborhood of the accepted grating space for that crystal. For NaCl a minimum occurs in the interval defined by $0.75a_0 < a < 1.25a_0$. It is quite likely that $w = -13590$ at $a = a_0$ is an actual minimum for NaCl, as it should be from theoretical considerations.

For NaF a minimum value of w lies in the interval $0.75a_0 < a < 1.50a_0$ but it appears to be closer to $a = 1.25a_0$ than to $a = a_0$. For LiF a minimum lies in the interval $0.75a_0 < a < 2a_0$ but it appears to be closer to $1.5a_0$ than to a_0. The reason for this shifting of the minimum value of w toward values of a greater than a_0 in the case of NaF and LiF is not clear. It may be that

[8] M. Born, 158–162 (1926).

[9] The term $1/2\sum_{k=1}^{8} e_k V_{k+}^1$ has been omitted for LiF.

w has not been determined accurately enough to make the difference between the values of w at a_0 and at $1.25a_0$ or $1.5a_0$ significant. Or it may be that we are not justified in assuming that a negative space charge can be replaced by an electron distribution in the case of NaF and LiF, since the total number of electrons per cell diminishes rapidly as we go from NaCl to LiF. NaCl has 112 electrons per cell, NaF has 80, and LiF has only 48.

Although we have determined approximately the position of a single minimum value of Φ for each of the three crystals, there remains the question as to whether or not this is the only minimum value of Φ. From the form of Eq. 10 and from the general relation between the structure factors and the grating space we hazard the guess that it is, although this conjecture has not been proved.

By the general method outlined in this paper one ought to be able to determine other crystal parameters besides grating spaces. Also it may be possible to relate Φ, the lattice energy which is not a directly measurable quantity, to other quantities which are measurable by some kind of a cyclical process as has been done by Born and others.

ORIGIN OF THE VARIATIONS IN THE SUN'S ROTATIONAL VELOCITY

By Ross Gunn
Naval Research Laboratory, Washington, D. C.

(Received August 23, 1930)

Abstract

The observed variations of the sun's rotation with solar activity are shown to be accounted for by changes of the electric and magnetic fields in the solar atmosphere. It is assumed that the sun's electric field arises from an electron or negative ion current flowing away from the sun's surface. The ionized atmosphere offers electrical resistance to the current and calculations show that the heat generated in the solar atmosphere is an appreciable fraction of the total radiated light energy. The electric field is thus correlated with the solar constant and since the solar constant is known to vary with the sunspot cycle, the atmospheric electric field, and hence the rotational period, must go through similar variations in magnitude. The subatomic processes which might give rise to the solar current are mentioned.

SIMULTANEOUS spectroscopic determinations of the sun's rotational velocity made at different observatories agree exceptionally well in magnitude but observations over long periods of time show that the velocities are not constant. The observed variations cannot be attributed to errors of observation or to a real change in the rotation of the entire solar mass. Newall[1] and Halm[2] have studied the observational data and definitely conclude that at times of maximum sun spot activity the sun revolves with an apparent speed that is greater by a few percent than the average; while at minimum activity, it rotates slower than the average. In addition to this type of variation there are observed temporary and local fluctuations which may amount to as much as ten percent of total superimposed velocity. These variations of velocity have been attributed to many causes, but no generally acceptable explanation has been given.

In an earlier paper[3] it was shown that the anomalies of the sun's rotation arose from the combined electric and magnetic forces acting on the ions of the solar atmosphere and that the only fundamental solar rotational period which can be observed directly, is the period of rotation of the sun's magnetic pole. With a knowledge of this fundamental period the surface velocity of the sun proper is readily calculated for any given point. To this velocity must be added the atmospheric drift which arises from the sun's magnetic and electric fields and may amount to as much as one quarter of the surface velocity. The magnitude of this superposed atmospheric velocity $\boldsymbol{u}$ is given by

$$\boldsymbol{u} = \frac{\boldsymbol{E} \times \boldsymbol{B}}{B^2} = \frac{E}{B} \sin \beta \qquad (1)$$

[1] Newall, Monthly Notices, R.A.S. **82,** 101 (1921).
[2] Halm, Monthly Notices, R.A.S. **82,** 479 (1922).

where E and B are the electric and magnetic fields respectively, and β is the angle between E and B. We have seen that this additional velocity has the correct value and distribution to account for the anomalous rotation if the sun carries a negative charge sufficiently great to produce a radial electric field of 0.013 volts/cm at a level in the reversing layer where the magnetic field is 25 gauss.

At any given latitude and level the angle β between E and B may be considered constant and it is clear from Eq. (1) that fluctuations in the velocity of the solar atmosphere are to be expected if the ratio of E to B changes for any reason whatsoever. Observations show that the atmospheric drift at any given level and latitude may increase by as much as ten percent from a time of minimum to maximum solar activity. We may attribute this increase in the atmospheric drift velocity to: (a) an increase in E; (b) a decrease of B; and (c) a readjustment of both E and B.

The magnetic field B at any given level in the solar atmosphere is determined primarily by the magnetic permeability of adjacent layers, and we have seen[3] that due to the diamagnetic action of the atmospheric ions the effective magnetic permeability is much less than unity. The intensity of magnetization I for the region is readily calculated and it is found[3] that

$$I = -\frac{NkT}{B(1 + (R/\lambda)^2)} \tag{2}$$

where N is the number of ions per cm^3, k the Boltzman constant, T the absolute temperature, λ the ion mean free path, and B the magnetic induction. The radius R of the circle generated by the ion as it spirals about the impressed magnetic field is given by

$$R = mV/Be = (2mkT)^{1/2}/Be \tag{3}$$

where m is the mass of the ion, V its velocity and e the ionic charge in e.m.u. Making use of the usual relations it is found that the magnetic permeability μ of the solar atmosphere at any level in terms of the observed magnetic field B is given by

$$\mu = \left[1 + \frac{4\pi NkT}{B^2(1 + (R/\lambda)^2)}\right]^{-1} = \left[1 + \frac{4\pi NkT}{B^2 + 2mkT/e^2\lambda^2}\right]^{-1} \tag{4}$$

In the reversing layer observation requires a permeability much less than unity and a long ion free path. Hence it follows from Eq. (4) that $4\pi NkT/B^2$ must have a value of the order of unity. This is the approximation which has been used heretofore to determine the pressure distribution in the solar atmosphere and it was found[3] that this approximation led to a mean atomic weight for the ions constituting the solar atmosphere of 3.3, a value in fair accord with the more recent unpublished determination of Professor Menzel.[4] The important thing to note in present discussion is that

[3] Gunn, Phys. Rev. **35,** 635 (1930); **33,** 614 (1929); **32,** 133 (1928); **34,** 335 (1929).

[4] H. N. Russell, Astro. Jour. **70,** 1 (1929).

the magnetic permeability of the solar atmosphere decreases with increasing values of (N) the specific ionization.

It seems well established that the value of the solar constant[5] increases by three percent from sunspot minima to maxima and that the radiation in the ultraviolet may increase by as much as thirty percent. This increase in radiation increases the specific ionization and decreases the permeability which in turn reduces the magnetic field at a given level. This reduction of the magnetic field increases the velocity of the atmospheric drift and at sun spot maxima we find, in accord with observation, that the solar rotational velocity is a maximum. It seems unlikely that this effect alone will account for the entire observed changes even though it is of the correct sign. We shall see that variations of the electric field appear to be of more importance in causing the observed changes in the rotational velocity.

In an earlier paper on the anomalous solar rotation[3] it was shown that the observed atmospheric motions demanded the existence of an electric field in the reversing layer which is directed radially inward and which has an equatorial value of 0.013 volts/cm at a representative level in the reversing layer. No attempt was made to account for the presence of this field, although the calculations showed that the potential difference between the sun and free space amounted to at least 10^6 volts and the phenomena giving rise to the electric field must therefore be quite energetic. The required electric field may arise from some unknown cause but at the present time it seems correct to assume that the presence of the electric field is evidence of a current of electrons or negative ions which flow constantly outward from the sun to free space. In order to calculate the total electric current discharged by the solar atmosphere on our assumption it is necessary to note that diamagnetism confines the entire solar magnetic field to the solar atmosphere, and ions which eventually leave the sun, no matter where they leave the surface, must drift across the magnetic field and not along it. We have seen[3] that the electrical conductivity of long free path ions moving across a magnetic field was very much smaller than the conductivity for ions moving in zero field or moving parallel to the magnetic field. In the present case we will not be greatly in error by taking the mean conductivity nearly equal to the conductivity of a typical level in the reversing layer near the equator. In this region the ions necessarily move across the magnetic field, and the usual expression for the electrical conductivity is reduced by a large factor dependent upon the ratio of the mean free path to the radius R of the ion path. If we take ion pressures in the reversing layer which are consistent with earlier work it seems likely that the electrons within the layer contribute but slightly to the electrical conductivity[3] and that the current is carried almost entirely by ions. The electrical conductivity (σ) of highly ionized stellar matter has been worked out by Chapman[6] who gives a mean value for the conductivity in the reversing layer of 1.7×10^{-11} e.m.u. This value is probably too large at the specific level in which we are interested, for here the magnetic field appreciably

[5] Report Astro. Observ. Smithsonian for 1923, p. 109.

[6] Chapman, Monthly Notices R.A.S. **89,** 54 (1928).

affects the conductivity of the ions as well as the electrons. We can do no better at the present time than adopt the foregoing value although it may be considerably in error. The mean current density i of negative charge flowing away from the sun is simply $i=\sigma E$ where $\sigma=1.7\times10^{-11}$ e.m.u. and E is 1.3×10^{6} e.m.u. Thus the current density in this region is roughly 2×10^{-5} e.m.u. per cm^2 and if we assume this special region to be representative of the whole surface, the total solar current amounts to 1.3×10^{18} e.m.u.

The potential difference between free space and the base of the reversing layer has been calculated[3] and found to approximate 1.5×10^{14} e.m.u. Moreover an extrapolation has been made extending down to the solar surface proper and the potential difference between the surface and free space was found to be 6.6×10^{15} e.m.u. This value may be taken as an upper limit and is based on the rather doubtful assumption that the drift velocities of the deep atmospheric layers are identical with the drift velocities in the reversing layer.

The solar atmosphere offers electrical resistance to the outward flow of current and the heat generated in the reversing layer by this process is radiated away into free space. The total electrical power dissipated in the form of heat in the solar atmosphere is simply the product of the total current and the potential difference. Thus the total electrical power dissipated in the reversing layer is 2×10^{32} ergs/sec and the total electrical power dissipated in all the surface layers cannot exceed 8×10^{33} ergs/sec. Since the total solar radiation is 3.8×10^{33} ergs/sec, we are driven to conclude that electrical heating effects in the solar atmosphere regulate to a certain degree the amount of light energy radiated into free space. The calculations which have been made are not sufficiently precise to warrant the conclusion that electrical heating effects are the only effects of major importance but they do indicate clearly that electrical phenomena may control to a marked degree the amount of energy transferred from the sun's interior to the solar atmosphere.

The introduction of an electrical mechanism to account for the transfer of energy from the interior of a star to its surface and hence, by radiation, to free space would profoundly modify many present astrophysical conceptions. It seems likely that the electrical effects will account for certain puzzling discrepancies which are now found in a comparison of the physical properties of different stars. A discussion of these astrophysical consequences must be postponed.

As a direct result of our conclusions in regard to the electrical energy dissipated in the solar atmosphere, we see that a direct relation exists between the amount of light energy radiated by the sun and the electric field produced in the reversing layer. Thus if the solar atmospheric current is increased by some subatomic or other process the radial electric field will increase, as will the total radiation. Observation shows[5] that at sun spot maxima the solar constant increases by a few percent, which indicates that the solar electrical field must have increased. The resulting increase in the electric field increases the superposed drift velocity of the solar atmosphere and this in turn increases the apparent observed rotational velocity. The above theoretical

discussion suggests that the electric field and hence the superposed drift velocity is approximately a linear function of the solar radiation, if we consider small variations. This is in complete accord with observation and is reliable evidence for the physical reality of the theory.

We have no direct clue to the phenomena which take place inside the sun and give rise to the outward negative atmospheric current. It is certain that similar phenomena give rise to the observed atmospheric current of the earth and because the potential differences encountered are so great it seems likely that the current systems arise from some subatomic transformation. G. C. Simpson[7] has suggested that electricity might be generated spontaneously within the earth while Swann[8] has considered the effects arising from the death of positive electricity. Swann did not discuss the important question of the origin of the energy for such a transformation. Unless the law of conservation of energy is entirely ignored, the death of electricity must be accompanied by the absorption of energy during its transformation, for death of the charge is electrically equivalent to its transfer to infinity. This requires energy and unless this energy is supplied by subatomic processes it seems clear that death cannot take place. W. Anderson[9] has avoided the above difficulties and he accounts for the earth's electric field by a single postulate. Anderson notes that the solar radiation must arise in some manner from the annihilation of matter and instead of assuming that the proton and electron vanish simultaneously with resultant radiation, or that the charge alone vanishes, he assumes that the proton's mass and charge both vanish while the associated electron wanders off into free space giving rise to certain solar and terrestrial phenomena. By this postulate Anderson was able to calculate the total solar current from the total solar radiation since one electron must appear for every proton annihilated and he found the total current to be 4×10^{16} e.m.u. Comparison of this value with 130×10^{16} e.m.u. computed above suggests that Anderson's postulate may yield the correct results, for the numerical value of the conductivity which has been chosen here, is probably too great. This postulate also connects directly the solar radiation with the solar electric current and therefore changes in the total radiation should be accompanied by rotational changes.

In his discussion of the source of solar energy Jeans[10] suggests that the energy arises from purely radioactive transformations of a material having a higher atomic weight than uranium. If this is the case it is not unlikely that the transformation results in a residue of electrons and perhaps several electrons may be released by the conversion to radiation of the mass of a single equivalent proton. By incorporating such hypothesis in the author's theory, the above numerical values can be brought into accord with solar radiation data. So far as can be determined at the present time, the author's postulate, or that of Anderson, cannot be shown to be inconsistent with any

[7] G. C. Simpson, Monthly Weather Review **44,** 121 (1916).
[8] Swann, Jour. Franklin Inst. **201,** 143 (1926).
[9] W. Anderson, Zeits. f. Physik, **42,** 475 (1927).
[10] Jeans, Astronomy and Cosmogony.

known fact. These postulates do, of course, readily account for the earth's observed electric field as well as that of the sun.

The anomalies of the solar rotation and its fluctuations, in the light of this and previous work, must be attributed primarily to the solar magnetic field and some special mechanism which gives rise to moderately large electric fields in the solar atmosphere. The electric fields are not so large that they could be detected by their Stark effect, and it is important to investigate carefully the less direct proofs of their existence. The theory is noncommittal in regard to the origin of the electric fields but the rough agreement between the present calculations and the values given by very simple and not unreasonable postulates are suggestive of some mechanism whereby subatomic transformations result necessarily in the release of negative electricity.

TRIBOTHERMOLUMINESCENCE

By R. E. Nyswander and Byron E. Cohn
Department of Physics, University of Denver

(Received August 22, 1930)

Abstract

Glass, certain crystals and frits are rendered thermoluminescent by the process of grinding, emitting light of very low intensity. The phenomenon is observed not only in samples of low purity but also in compounds prepared from chemically pure grade materials. Quantitative measurements of light intensity were made by means of a polarization photometer. The particles tested were graded in size by means of a set of screens ranging from 20 to 200 mesh. The results show that the quantity of emitted light depends upon the nature of the substance and the size of the particles, and diminishes slowly with the time after grinding. Common glass tubing crushed to various grain sizes emitted the maximum amount of light when the linear dimensions of the average mesh size fell between 0.015 cm and 0.025 cm respectively.

CERTAIN substances, when powdered, store up energy which is released by heat in the form of light of very low intensity. As in other types of thermoluminescence, the temperature of emission is below that of incandescence. Wiedemann[1] applied the term triboluminescence to the light given off when certain materials are rubbed or scratched and Levison[2] designated as tribophosphorescence the luminescence of more prolonged duration excited by rubbing. The term tribothermoluminescence has been given to the phenomenon under consideration which involves the process of grinding followed by the application of heat.

Tribothermoluminescence has been observed in crystals, glasses and frits. In all cases the material tested was first heated in an electric furnace to destroy any trace of thermoluminescence which might have been present, and then ground in a mortar and screened to the desired size. Tribothermoluminescence was observed in clear, colorless crystalline fluorite, and also in lithium borate which forms a white vitreous opaque mass (frit) when fused, as well as all samples of clear colorless glass tested with the exception of a specimen of chemically pure boric anhydride.

The energy of triboluminescence has been discussed by Imhof,[3] Trautz[4] and Schmidt.[5] One of these suggested explanations involves the recombination of electrons displaced by the process of rubbing, in the other the energy is the result of small electric sparks induced by differences of potential, the electrical doublet theory. In the case of tribothermoluminescence the latter

[1] E. Wiedemann, Wied. Ann. **34,** 446 (1888); Nature **27,** 181 (1889).
[2] W. Levison, Science **19,** 826 (1904).
[3] A. Imhof, Phys. Zeits. **18,** 78 (1917) and **20,** 131 (1919).
[4] M. Trautz, Ion **2,** 77.
[5] H. Schmidt, Phys. Zeits. **19,** 399 (1918).

theory does not seem directly applicable after the following test. The particles immediately after grinding were immersed in an electrolyte of dilute nitric acid which would allow the recombination of all electric charges which might result from the process of grinding. The glass was then washed with distilled water filtered and dried in the dark. After this treatment to remove all electric charges, the glass showed the usual thermoluminescence.

Ultraviolet light is very active in exciting the energy of the thermoluminescence in glass.[6] If ultraviolet light is emitted in the process of grinding, this might be the source of energy of the thermoluminescence of the glass particles. Accordingly, three similar samples of glass were tested in the following manner. The first was ground in a mortar in the usual way, i.e. in the presence of air. The second sample was ground under water then washed in distilled water and dried. The third sample was ground in a dark solution washed and dried as before. The presence of the water and the opaque solution would re-

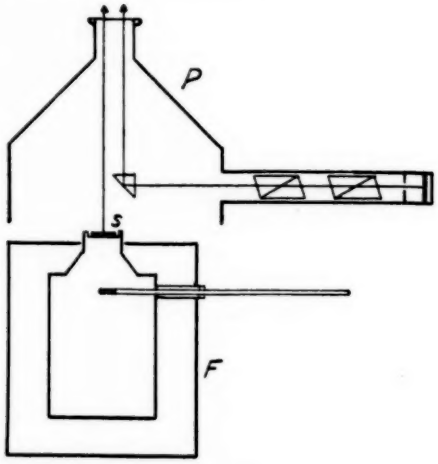


Fig. 1. Diagram of apparatus showing arrangement of polarization photometer and furnace.

duce through absorption the light radiated to the adjacent particles by the process of grinding. When these samples were tested for thermoluminescence the first emitted a much larger quantity of light than the second, and the second emitted more than the third, which gave out scarcely more than the threshold value.

All light measurements were made by means of a polarization photometer.[7] The usual form was modified to allow a larger angular contrast background, as shown in Fig. 1, in which *P* is the photometer, F the furnace and S the substance to be tested. The samples were heated in small brass cups with cavities 0.8 cm diameter and 0.4 cm deep, which were placed on top of the large brass post of the electric furnace. By means of a thermometer in the furnace it was possible to hold the temperature of the sample constant for any period. Light measurements were made at time intervals from the

[6] Nyswander and Cohn, J.O.S.A. **20,** 131 (1930).

[7] Nyswander and Lind, J.O.S.A., and R.S.I. **13,** 651 (1926).

beginning to the end of the period of luminescence. These light intensities were plotted as ordinates with the corresponding time intervals measured from the beginning of luminescence as abscissas. The total quantity of light emitted by the sample is proportional to the area under the curve. These areas were measured by means of a planimeter. All light measurements are expressed in terms of relative values.

The quantity of light emitted by the material was found to be dependent upon (1) the nature of the substance, (2) the size of the particles, (3) the time elapsed after the process of grinding. The effect of the nature of the material is brought out by the following measurements of the light given out by fused sodium borate glass as compared with zinc borate glass (45.5 percent zinc

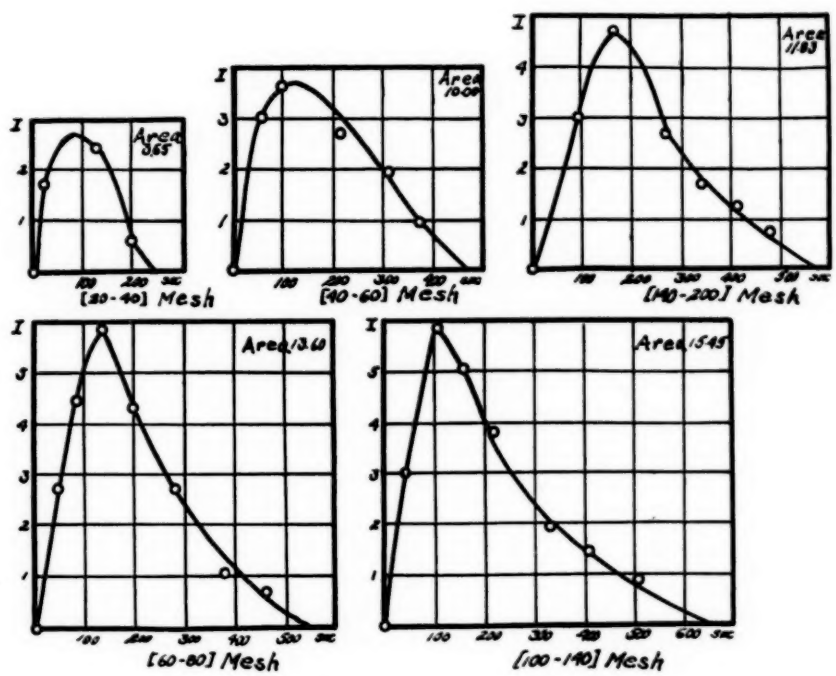


Fig. 2. Total quantities of light emitted by glass for various screen sizes.

oxide to 54.5 percent boric acid). The total quantities of emitted light in the two cases are 16.5 and 1.72 respectively. The two sets of measurements were made under similar conditions. The particles were sieved between 180 and 200 mesh screens, and the temperature of the furnace was 330°C. The results also show that chemically pure grade materials sometimes exhibit tribothermoluminescence.

To detemine the effect of grain size on the quantity of light emitted, the particles were separated into groups by means of a series of screens. To obtain the size of the screen openings, eight measurements were made on different parts of each screen by use of a comparator, and the mean value used. Ordinary soft glass tubing was heated to remove all traces of luminescence

and then ground in a porcelain mortar and sieved. The graded sizes were heated in the furnace held at 360°C and a series of intensity-time curves obtained and plotted as shown in Fig. 2. The areas of these curves were then used as ordinates with the mean of the two corresponding mesh openings as abscissas and plotted as shown in Fig. 3. Curve *A* represents the total light energy for the graded sizes two days after powdering, while *B* was determined 24 days after grinding. The difference between these two curves clearly

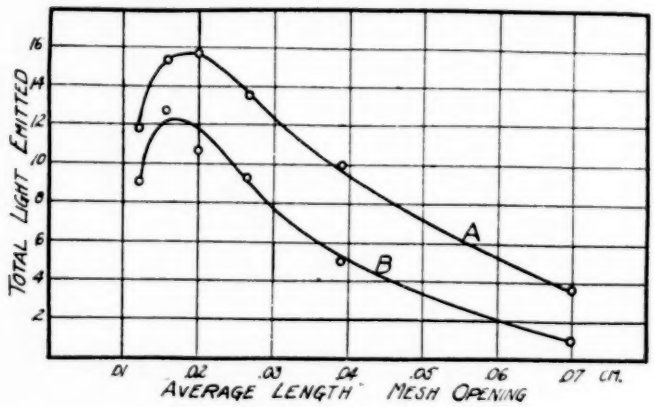


Fig. 3. Variation of emitted light as dependent upon grain size.

shows the decay of luminescence with time. The smaller the size of the particles the greater the amount of energy stored, however, the smaller the size of the particles the less the quantity of light which reaches the surface through a given thickness of the powdered material. These two factors would seem to determine the position of maximum light emission of the curves *A* and *B*. Tribothermoluminescence has been observed in glass of sizes less than 100 mesh one year after it had been powdered.

LETTERS TO THE EDITOR

Prompt publication of brief reports of important discoveries in physics may be secured by addressing them to this department. Closing dates for this department are, for the first issue of the month, the twenty-eighth of the preceding month; for the second issue, the thirteenth of the month. The Board of Editors does not hold itself responsible for the opinions expressed by the correspondents.

High-Voltage Tubes

As a more detailed paper will perhaps not be justified for some time, it may be of interest to report that we have obtained measurements which satisfactorily verify the production of high-speed electrons and of very penetrating x-rays, in the region above 1,000 kilovolts, using the high-voltage Tesla coils and cascade tubes previously described,[1] and that the $H\rho$-values of the highest-speed electrons roughly correspond in voltage equivalents (according to the standard β-ray tables) to the peak voltages of the Tesla coil as measured by the capacity-potentiometer. These measurements have been carried out at voltages considerably below the maximum voltages which we have previously applied to tubes, to reduce the insulation difficulties which are experienced when working near the limit which the apparatus will attain.

A 12-section tube placed inside of a 6-inch diameter Tesla coil having taps on the winding connected to the separate electrodes of the tube has been used, the grounded end of the tube projecting through the oil-tank between the poles of an electromagnet. A slit-system defines a beam of electrons which strikes at will a fluorescent screen or a photographic plate enclosed in 3, 6, and 10 sheets of 0.01 mm aluminum-foil. The distribution of the magnetic field is carefully determined by the triangular search-coil method described in Thomson's "Positive Rays". Unfortunately only very rough spark gap voltage-measurements were obtained during these tests, but the indicated voltages for a series of four plates, for example, were about 900 kilovolts for the first two and 1,000 to 1,300 kilovolts for the last two, the highest $H\rho$-values for which electrons were definitely detected on the plates corresponding to 850, 980, 1,110, and 1,130 kilovolts respectively. The intensity of the electron-stream at the plate is rather low, approximately 500 primary sparks being required for a satisfactory exposure-density through 0.10 mm of aluminum at the maximum $H\rho$-value. A 2-mil tungsten-filament constituted the electron-source at the high-voltage end of the tube; with the filament cold the tube was perfectly dark and no electrons could be detected.

A Geiger-Müller tube-counter mounted in a 6-inch lead-shield at a distance of 55 cm to one side of the copper slit-system in the tube showed no evidence of radiation through 1 inch of lead when the tube was operated as above described. Due to the very short on-time of the high-voltage pulses (resonant frequency about 30,000 cycles) and the very small part of the emission from the filament which would thus be effective, only a very small intensity of radiation could be expected, and consequently a "flashing" tube was resorted to in order to have a more copious emission available at the peak of the voltage-wave. The first observation showing penetrating radiation was made in January, using a similar "flashing" tube. The tube and the Tesla coil were placed end to end in the tank, thus applying the high voltage directly to the end electrodes only, causing it to flash; all other electrodes were electrically floating (no potentiometer was used, but external shields surrounded each tube-section as previously described).[1] This tube was operated at a voltage of 1,100 to 1,300 kilovolts, the $H\rho$ of the electrons being continuously observed on the fluorescent screen. As the voltage required to cause the tube to flash slowly increased, and the intensity of the flashing was somewhat irregular, absorption measurements were necessarily rough, but

[1] Phys. Rev. **35**, 51, 66, and 1406 (1930).

during counting periods of about ten minutes, running the Tesla coil at 65 sparks per minute, approximately six extra counts per minute were produced through a 1-inch lead absorption-screen, two extra counts per minute through 2 inches of lead, and 28 extra counts per minute through 5/8 inch of lead. The absorption coefficients obtained from the above data are thus of the expected magnitude. The sensitivity of the Geiger counter, which has a residual of about 20 counts per minute, is such that 0.105 mg of radium at a distance of 75 cm produces 26 extra counts per minute through two inches of lead (89 extra counts per minute through 1 inch of lead). The average intensity of the radiation from the tube (at only 66 sparks per minute) is thus very low, but the estimated instantaneous radium equivalent is considerable, and this can be increased by the use of a tube adapted for the production of radiation, the above data being obtained with the tube arranged for the $H\rho$-measurements.

The production of "artificial β- and γ-rays" in the region above 1,000 kilovolts has thus been demonstrated. Work is now in progress on the acceleration of protons to radioactive speeds.

M. A. Tuve
L. R. Hafstad
O. Dahl
Department of Terrestrial Magnetism,
Carnegie Institution of Washington,
September 11, 1930.

Note added September 12, 1930.

By an amusing coincidence the August 29 issue of Die Naturwissenschaften arrived just after the writing of the above note, and it contains a letter by A. Brasch and F. Lange describing the spectacular performance of their high-voltage tube at 2.4 million volts. They are to be congratulated without reserve on the success of their work.

M. A. T., L. R. H., O. D.

Use of the Pierce Acoustic Interferometer for the Determination of Absorption in Gases for High Frequency Sound Waves

In recent articles in this journal, W. H. Pielemeier (Phys. Rev. **34,** 1184, 1929: Phys. Rev. **36,** 1005, 1930) reports determinations of the absorption coefficient in gases for high frequency sound. He employs both the Pierce acoustic interferometer and a torsion vane method. The observations by the first method reveal that as the sound path in the gas is increased the changes in plate current through the tube diminish, and from this rate of diminution Pielemeier attempts to calculate the absorption coefficient. We do not believe this procedure to be entirely justified. The variations in plate current are a function not only of the absorption in the gas but of the circuit constants as well. For we are here dealing with decrements in coupled circuits. The network may be considered as made up of a crystal circuit with a decrement which depends upon the reflector position in the gas; also this crystal is coupled to a plate circuit with its own decrement. Hence, uncorrected observations on changes in plate current through the tube give no information concerning the absolute absorption in the gas. Experiment has shown us that these changes in current depend both on the reactance and the decrement of plate circuit.

Elias Klein
W. D. Hershberger
Naval Research
Laboratory,
Bellevue, D. C.
September 15, 1930.

The Magnetic Moment of the Li_7 Nucleus

Using the experimental data on the hyperfine structure of the alkalis, Fermi[1] and Hargreaves have computed approximate values of the magnetic moments of their nucleii. Goudsmit and Young[2] have estimated the magnetic moment of the Li_7 nucleus on the simplifying assumption that the coupling to the nucleus takes place entirely through the 1*s* electron. The apparent inconsistency of their results may be due in part to the experimental difficulty of separating the structure of the $^3S_1-P_1$ and $^3S_1-P_2$ groups. In part it is probably due to their using $(1/2)(h/2\pi)$ for the nuclear

[1] E. Fermi, Zeits f. Physik **60,** 320 (1930). J. Hargreaves, Roy. Soc. Proc. **124,** 568 (1929); **127,** 141 and 407, (1930).

[2] S. Goudsmit and L. A. Young, Nature **125,** 461 (1930).

spin, their main object being to show that the magnetic moment is of the order of magnitude of the theoretical magnetic moment of the proton.

It has since been shown by Granath[3] that the main reason for using $i=1/2$ for the nuclear spin is absent. The frequency separations between components agree best with $i=3/2$ which is also in agreement with the band spectrum results of Harvey and Jenkins.[4] Goudsmit, Güttinger[5] and ourselves have known for some time that this value of i is fairly consistent with the experimental pattern. The nuclear g factor is on this hypothesis in the neighborhood of 2.3, this value being obtainable directly from Schüler's frequency table and the formula used by Goudsmit and Young.

We have undertaken to make a more accurate calculation for the g factor. Our special reason for doing so is that the electronic configuration of the $^3S_1$ level is especially simple and appears to be favorable for accurate results. We have also been intrigued by the possibility that the g factor might be 2 exactly which would speak in favor of regarding the magnetic moment of Li_7 as due to three protons acting independently. It is our purpose to report here briefly on the results of our calculations.

The work of Casimir and of Fermi shows in the case of one electron that it is essential to consider the problem from the point of view of Dirac's electron equation. A non-relativistic equation cannot be expected to give even approximately correct results for S electrons because in the immediate vicinity of the nucleus the electron velocity cannot possibly be regarded as small. There exists at present no satisfactory relativistic treatment of two particles. Nevertheless, for the present purpose, it is possible to form a reasonable extension of the one-electron treatment. In a discussion of light atoms it is possible to treat the one-electron problem also by means of an equation of the Pauli-Darwin type, i.e. employing two rather than four components for ψ. Darwin[6] has shown how such an equation can be derived from the original four-component Dirac form. The interaction energy with the nuclear magnetic moment μ is then to a sufficient approximation

$$\begin{aligned}H' &= (e/mc)\left[1+(E-mc^2\right.\\ &\quad \left.+eA_0)/2mc^2\right](\boldsymbol{M}\boldsymbol{\mu})r^{-3}\\ &\quad +(hec/2\pi)\left[2mc^2+eA_0\right]^{-1}\{-r^{-3}(\boldsymbol{\mu}\boldsymbol{\sigma})\\ &\quad +3(\boldsymbol{r}\boldsymbol{\sigma})(\boldsymbol{r}\boldsymbol{\mu})\}\\ &\quad +(hec/2\pi)\left[2mc^2+eA_0\right]^{-2}(e\mathcal{E})\{r^{-2}(\boldsymbol{\mu}\boldsymbol{\sigma})\\ &\quad -r^{-4}(\boldsymbol{r}\boldsymbol{\sigma})(\boldsymbol{r}\boldsymbol{\mu})\}\end{aligned} \tag{1}$$

The charge of the electron is taken to be $-e$, E is the total energy, m and c are respectively the mass and the velocity of light, A_0 is the electrostatic potential due to the nucleus, $\mathcal{E}$ is the electric intensity due to the nucleus, $\boldsymbol{r}$ is the displacement vector from the nucleus to the electron, $\boldsymbol{M}$ is the angular momentum vector operator, and σ is Pauli's spin vector. In the first term the square bracket can be omitted for anything but s terms. For s terms it enforces the convergence of the otherwise divergent integrals for r^{-3}. The second term is the closest analogy of the dipole interaction of the nuclear and electronic magnetic moments. On account of the square bracket its effect disappears for s terms. The third term, involving the electric intensity $\mathcal{E}$ is negligible except for s terms. If the square bracket in the second term were absent the result would be indeterminate. In fact it may be shown that if μ were the limit of a space distribution of magnetization the result of the second term would be $(-1/2)$ of the correct total result.

For two electrons we postulate the interaction energy to be the sum of two terms of the type of (1) one term in each electron. The first order perturbations for the energy of a ^3S term come out without further approximations to be given by

$$\begin{aligned}w &= w_1(1,-(1/i),-(i+1)/i)\\ w_1 &= (16\pi/3)\mu\mu_0\textstyle\int\psi^2(0,q)dq\end{aligned} \tag{2}$$

where μ_0 is the Bohr magneton and $\psi(q_1, q_2)$ is the nonrelativistic Schroedinger wave equation. The letter q_1 stands collectively for the Cartesian coordinates of electron 1, and simi-

[3] L. P. Granath, Phys. Rev. **36,** 1018, (1930).

[4] A. Harvey and F. A. Jenkins, Phys. Rev. **58,** 789, (1930).

[5] We should like to express at this point our appreciation to Professor Goudsmit and to Professor Pauli. It is through the courtesy of the latter that we have received proofs of a paper by Güttinger which is in press in the Zeitschrift für Physik and which shows in detail how $i=3/2$ accounts for Schüler's pattern.

[6] C. G. Darwin, Proc. Roy. Soc. **118,** 654, (1928).

larly q_2 for those of 2. The splitting is proportional to the probability of finding an electron in a unit volume at the nucleus. By means of (2) one can calculate a correction factor to be used in the formulas of Goudsmit and Bacher.[7] This factor is

$$f=2\int\psi^2(0, q)dq/\int\psi_1 S^2(q)dq.$$

In order to obtain f and hence μ it is necessary to find a solution of the two-electron non-relativistic wave equation. In the absence of an analytic solution we have tried a number of wave functions adjusted for the minimum energy by the variational method. Among these the function

$$(r_1-c)\exp\left[-(a/2)r_1-(b/2)r_2\right] -(r_2-c)\exp\left[-(a/2)r_2-(b/2)r_1\right]$$

appears to be best suited for the purpose. It gives when minimized for a, b, c an Eigenwert which is in excellent agreement with experiment. In terms of the ionization potential of Li^{++} the Eigenwert obtained is -1.1354 which may be compared with the spectroscopic value -1.1358. The best values of a, b, c are 0.38, 1.00, 3.18. The correction factor f is for these values of a, b, c, $f=1.063$. Using this factor and making the slight correction in the frequency of Schüler's component,[3] mentioned by Granath,[3] the g factor becomes 2.13 for $i=3/2$. This is closer to 2 than the uncorrected value 2.3. The remaining difference of 6% appears to be too large to be accounted for by experimental error. The accuracy of the theoretical calculation may, of course, be questioned. An estimate of it may be made by the method recently published by Eckart.[8] Such an estimate would lead one to suppose that the result is accurate only to about 7%. In this special case, however, our experience with other trial functions as well as a more detailed consideration of the possible errors indicate that Eckart's accuracy criterion is likely to be too conservative. *The present evidence is, therefore, that* if the nuclear *spin is 3/2, Schüler's observed frequency* separation between *components* (1) *and* (3) *speaks in favor of a nuclear magnetic moment greater than that of three protons by about* 6%.

It should be noted that we base ourselves entirely on Schüler's observed frequency separation between components (1) and (3) because the other components of the pattern are not sufficiently resolved to make definite conclusions possible. It is also not out of place to emphasize here that in our calculation the magnetic field due to the electrons has been supposed to have no effect on the nuclear magnetic moment. An estimate shows this field to be about 4.5×10^6 gauss. An exact experimental test of the interval ratio given by (2) may show to what extent the nuclear magnetic moment is affected by this magnetic field.

G. Breit
F. W. Doermann

Department of Physics,
New York University,
September 15, 1930.

[7] S. Goudsmit and R. F. Bacher, Phys. Rev. **34,** 1501 (1929).

[8] Carl Eckart, Phys. Rev. **36,** 878 (1930). Note also Eckart's function for ^3S closely resembling ours.

Wind Mixing and Diffusion in the Upper Atmosphere

I am glad to see that Professor Chapman (Phys. Rev. **36,** 1014, 1930) agrees with me in thinking that there is no question of Dr. Maris' priority in the development of the subject of wind mixing and diffusion in the upper atmosphere. At the same time Professor Chapman claims that his work was independent of that of Dr. Maris. The internal evidence in his paper (Proc. Roy. Soc. A, **122,** 369, 1929) does not support this claim. To give one example; on page 375 he wrote, "It is here assumed that the mixing ceases at 110 km," with no indication of any calculation which would justify such an exact assumption. Maris had already written (Nature, December 1927) "This 'diffusion' level for hydrogen would move from infinity down to 142 km in one day, at the end of five days it would be at a height of 127 km and in 50 days it would be at 113 km. The corresponding levels for helium would be at 137, 120 and 106 km, respectively." And in Terr. Mag. **33,** 233 (1928) after five printed pages of calculation and close physical reasoning Maris gave (Table 3) the diffusion levels of six atmospheric gases for summer, winter, day and night which "averaged about 110 km."

E. O. Hulburt

Naval Research Laboratory,
September 19, 1930.

BOOK REVIEW

Einführung in die Mechanik und Akustik. Dr. R. W. Pohl, Professor of Physics in the University of Göttingen. Pp. 240, figs. 440. Julius Springer, Berlin, 1930. Price R.M. 15.80.

This book is the first part of Professor Pohl's lectures concerning experimental physics. Other parts are to appear within a few years. This first part is not a complete text of mechanics and acoustics, but rather an exposition of the fundamental experiments in this subject. Mathematical demonstrations are generally omitted. The book contains numerous lecture experiments that are quite new to the reviewer. In its 244 pages there are 440 figures illustrating the experiments performed. In many of these shadow pictures are used instead of photographs of the apparatus. These seem to be a successful feature of the book. The author has prepared a work that will prove interesting and stimulating to students and teachers of physics and to others interested in the subject.

G. W. STEWART

Differential Equations, F. R. MOULTON

There is a large class of people, and among them most physicists, who have occasion to solve differential equations, who like to think with care, and who feel more or less ashamed of themselves when they proceed to do all sorts of curious things with series in seeking solutions of the equations in the doubtfully happy belief that somewhere or other there are some mathematicians or theorems so "pure" that they may shoulder all the iniquities of the mathematically wicked and somehow or other show that they are justified. Yet, when these lost souls turn to the realms of pure analysis for guidance, they are apt to be appalled by the necessity of wading through an infinite complexity of unappealing arguments and symbols—of η's which are smaller than ϵ's and can never be smaller than ξ's, and so forth. To all such as these, Professor Moulton's book should be of the greatest value, representing as it does, the viewpoint of one interested fundamentally in the logical foundations of the subject, and at the same time, in its practical applications in astronomy, ballistics, etc.

The book is concerned almost entirely with "ordinary" differential equations, a couple of pages on "partial" differential equations being introduced merely to show the general relation of the considerations there involved to those of ordinary differential equations. Professor Moulton commences by emphasizing the fact that any problem having as its outcome the solution of a set of differential equations may be considered, logically, in four parts—formulation of the problem in terms of the properties of the functions involved—proof of the existence of a solution of the problem—determination of the properties of the solution—development of practical means of obtaining the solution. Then follows a discussion of the reduction of differential equations to "Normal Form." The second chapter is devoted to the formal construction of solutions in terms of power series in the independent variable, and to the subject of convergence of the solution. Chapter III deals with the construction of solutions in the form of power series in a parameter, again with a careful discussion of the question of convergence. Chapter IV treats the subject of solution by the method of variation of parameters. The general procedures in all of these cases are amply illustrated by practical examples; and, in the form of an appendix will be found a very useful collection of such theorems in analysis as are used in the logical development of the text. The appendix should be of particular value to the physicist whose acquaintance with these theorems dates back to a not too complete comprehension of their content gained in his student days, and who would otherwise lose much time in digging them out of the standard treatises on analysis. After devoting Chapter V to Integrals of Differential Equations and Chapter VI to Analytic Implicit Functions, the author proceeds in Chapters VII to X, to take up the solution of special problems which, while of fundamental practical interest in themselves, are chosen primarily to illustrate different methods of setting about the solutions of problems, having in mind the conditions and requirements of the special problem in hand. In the chapter on the "Sine Amplitude Function" in particular, the author emphasizes

how it is possible to obtain much information regarding the nature of a function defined by a differential equation without recourse to an actual solution of the equation itself.

Chapters XI and XII contain an admirable account of the method of solution by "Successive Integrations" and one acquires a certain aesthetic satisfaction from the author's emphasis on the fact that the process is not fundamentally an approximation, but is, in the limit, logically exact in the same sense that other limiting processes in analysis such as the expression of functions in series, may be regarded as logically exact. In Chapter XII the application of the method is set out in detail, even to the extent of suggestions in the matter of actual numerical procedure.

After devoting Chapter XIII to the Cauchy-Lipschitz method of solution the author proceeds, in Chapters XIV to XVII to discuss in detail the subject of linear differentials of various types, with applications. The last chapter is devoted to "Differential Equations with Infinitely Many Variables."

The work throughout is written in an exceptionally clear vein. Even in the parts of a more purely mathematical nature, conglomerations of symbols are minimized in favor of the spirit of the processes discussed, so that the reader is brought to a sympathetic interest in the significance and reality of content of the steps taken in place of being left, as is so often the case, with profound respect for the complexities of the subject coupled with a sense of relief that he does not have to do that kind of thing himself.

The addition of Historical Notes at the end of each chapter serves further to humanize the subject, and, by a curious freak of the mind, tends to render more clear the purport of many things by sheer realization of the fact that these matters were not always clear to the pioneers of the science itself.

W. F. G. SWANN